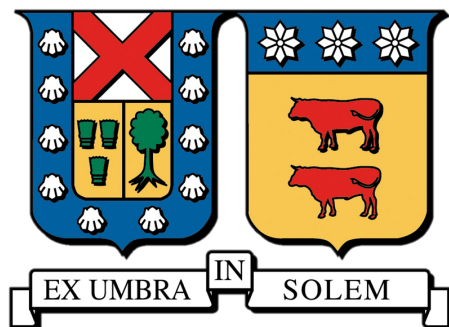




UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS

**EFFECTO DEL EMPAREJAMIENTO DE GÉNERO ENTRE DOCENTE Y
ESTUDIANTE EN EL RENDIMIENTO ACADÉMICO EN CHILE:
ANÁLISIS DE MECANISMOS SUBYACENTES**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

AUTORA

LAURA ESCOBAR SANJINES

PROFESOR GUÍA

Dr. BERNARDO EUGENIO PINCHEIRA SARMIENTO

PROFESORA CO-REFERENTE

Dra. PAULINA SOLEDAD SANTANDER ASTORGA

VALPARAÍSO, SEPTIEMBRE 2026



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Efecto del emparejamiento de género entre docente y estudiante en el rendimiento académico en Chile: Análisis de mecanismos subyacentes.

Nombre del candidato(a): Laura Escobar Sanjines

Carrera / Grado: Ingeniería Civil Industrial

Campus: Casa Central **Departamento:** Departamento de Industrias

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Bernardo Pincheira Sarmiento, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

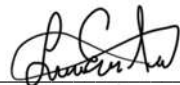
Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 16-09-2026

Firma: 

Estudiante o Candidato(a):

Fecha: 16-09-2026

Firma: 

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Tabla de Contenidos

1 Agradecimientos	5
2 Problema de Investigación	6
3 Objetivos	9
3.1 Objetivo General	9
3.2 Objetivos Específicos	9
4 Marco Teórico	11
4.1 Dinámicas de género en el aula	11
4.1.1 Mecanismos subyacentes: modelo de rol y sesgo docente	12
4.2 Contexto educativo chileno	14
4.3 Evaluación del rendimiento académico a través del SIMCE	16
4.4 Modelo econométrico	18
4.4.1 Ajuste mediante efectos fijos	20
4.5 Análisis e interpretación de los resultados con perspectiva de género	22
5 Consolidación de datos	25
5.1 Especificación de bases de datos utilizadas del SIMCE 2017	26
5.1.1 Base de alumnos	26
5.1.2 Cuestionarios docentes	27
5.1.3 Cuestionario de estudiantes	29
5.1.4 Cuestionario de apoderados	31

5.2	Construcción de indicadores	32
5.2.1	Indicadores de mentalidad	33
5.2.2	Indicadores asociados a Matemática	34
5.2.3	Índice de condiciones laborales docentes	35
5.3	Unión de bases y construcción de la base final	37
6	Estadísticas descriptivas	39
6.1	Caracterización general de la muestra	39
6.2	Descripción de las principales variables	40
6.3	Diferencias descriptivas según género	41
6.3.1	Prueba $t - test$	43
6.4	Características familiares	45
7	Metodología	47
7.1	Estimación del efecto del género docente sobre el rendimiento estudiantil	47
7.1.1	Modelo 1: control por rendimiento académico previo	48
7.1.2	Modelo 2: incorporación de características familiares	50
7.1.3	Modelo 3: incorporación de características docentes	51
7.1.4	Modelo 4: efectos fijos por establecimiento	53
7.1.5	Modelo 5: efectos fijos por estudiante	54
7.1.6	Modelo 6: efectos fijos por curso-asignatura	56
7.2	Estrategias para evaluar mecanismos subyacentes	57
7.2.1	Estrategias asociadas al sesgo docente	57
7.2.1.1	Expectativas académicas del docente	57

7.2.1.2	Composición de género del curso	59
7.2.2	Estrategias asociadas al modelo de rol	60
7.2.2.1	Diferencias entre asignaturas	61
7.2.2.2	Educación de la madre	62
7.2.2.3	Autoconcepto matemático	63
7.2.2.4	Mentalidad hacia Matemática	64
8	Resultados	66
8.1	Relación entre el género docente y el rendimiento estudiantil	66
8.1.1	Resultados con controles observables	66
8.1.2	Resultados con efectos fijos	69
8.2	Análisis de los mecanismos subyacentes	74
8.2.1	Evidencia asociada al sesgo docente	75
8.2.2	Evidencia asociada al modelo de rol	76
9	Conclusiones	80
10	Limitaciones	83
11	Anexos	85
11.1	Anexo A: Preguntas utilizadas en los indicadores de mentalidad	85
11.2	Anexo B: Pregunta utilizada para la mentalidad hacia Matemática	86
11.3	Anexo C: Ítems utilizados para el autoconcepto matemático	87
11.4	Anexo D: Variables utilizadas en el índice de condiciones laborales	88
11.5	Anexo E: Robustez a la categorización del ingreso del hogar	89



11.6	Anexo F: Resultados completos de los modelos con controles observables	91
11.7	Anexo G: Resultados completos de los modelos con efectos fijos	95
11.8	Anexo H: Pruebas complementarias de identificación	98
11.9	Anexo I: Resultados completos asociados a los mecanismos subyacentes	101

1 Agradecimientos

Este camino, que hoy llega a su fin, no habría sido lo mismo sin las personas que me acompañaron, por lo que quiero tomarme este espacio para agradecer, en primer lugar, a mi familia, especialmente a mi mamá, Lourdes, quien ha sido mi guía y apoyo en cada momento. Mamá, gracias por cada enseñanza, por tu esfuerzo y por haber buscado siempre las mejores oportunidades para mí; eres mi ejemplo de fortaleza y perseverancia, y todo lo que soy tiene una parte de ti. A mis ratas, Nirvana e Ian, gracias por acompañarme en esta etapa junto a mi pequeño Toby y por ser parte tan importante de mi vida; Nirvana, por estar a mi lado desde niñas y ser siempre ese apoyo en los momentos difíciles, e Ian, porque tu llegada nos unió muchísimo como familia y me enseñaste, desde tu propia historia, el valor de la resiliencia y de siempre encontrar motivos para sonreírle a la vida.

A mi mejor amigo, Tomás, quien aportó a mi mundo de intensidad esa cuota de racionalidad que tantas veces necesitaba, me escuchó y acompañó siempre, especialmente durante los momentos más difíciles de la U, y con quien guardo con muchísimo cariño tantos recuerdos. A mis tilines, Coni, Paula, Cegarrak y Vale, gracias por compartir conmigo risas, frustraciones, horas de estudio y carretes, por hacer chistosos los días difíciles y por no dejarme sola en ni un solo cumpleaños desde que los conocí. A mi nuevo personaje desbloqueado, Alan, mi pareja, quien me ha acompañado estos últimos meses, gracias por escucharme, apoyarme y llenar mis días de alegría. Me hace muy feliz coincidir contigo en este momento de mi vida y espero que podamos seguir disfrutando juntos de todo lo lindo que nos queda por vivir.

A mi profesor guía, Bernardo Pincheira, gracias por su preocupación, por estar presente durante este proceso y por su orientación y apoyo en el desarrollo de mi investigación.

A todos ustedes, gracias por su cariño, sus enseñanzas y por haber sido parte de esta etapa; cada uno dejó algo en mí que me llevaré para siempre.

2 Problema de Investigación

Desde hace décadas, una de las principales problemáticas que enfrenta la sociedad es el bajo nivel de participación de las mujeres en carreras STEM (Science, Technology, Engineering, and Mathematics), que, a pesar de los avances en igualdad de género, sigue siendo limitado, ya que las mujeres representan solo el 29,7 % de la matrícula de pregrado, el 29 % de los programas de magíster y el 35,9 % en doctorado, según la Cuarta Radiografía de Género en Ciencia, Tecnología, Conocimiento e Innovación (Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, 2025).

Las explicaciones de esta brecha pueden agruparse en tres categorías: la psicológica, centrada en factores individuales como la autoeficacia y las preferencias; la sociocultural, que resalta la influencia de la cultura, los estereotipos, la discriminación y los roles de género; y la biológica, que sostiene que las diferencias anatómicas entre sexos influyen en las elecciones profesionales. Dentro de las categorías psicológica y sociocultural, resulta relevante analizar cómo los estereotipos y roles de género presentes durante la trayectoria escolar pueden influir en el rendimiento académico y, en consecuencia, en la futura participación en carreras STEM. En línea con esto, también resulta necesario considerar la influencia de los distintos actores que forman parte de esta trayectoria, entre los cuales destaca el docente, cuya influencia sobre el comportamiento y desempeño del estudiante se complementa con la ejercida por la familia y los pares (Morales Inga y Morales Tristán, 2020).

Trasladando esta discusión al contexto chileno, Paredes (2014) estudia el impacto del emparejamiento de género entre profesor y estudiante en el rendimiento académico, utilizando principalmente datos SIMCE de 2009 correspondientes a estudiantes de octavo básico. La autora demuestra que las niñas se benefician al ser asignadas a profesoras, mientras que los niños no experimentan efectos negativos, y señala que la evidencia obtenida sugiere que este resultado sería más consistente con un mecanismo de modelo de rol que con uno asociado al sesgo docente.

No obstante, debido a que los datos son previos a diversas reformas educacionales y tienen más de diez años de antigüedad, es probable que los patrones de emparejamiento y los efectos estimados no reflejen el contexto actual. Por lo tanto, resulta necesario actualizar el análisis del impacto del género docente en el rendimiento estudiantil, enfocándose en los mecanismos subyacentes de modelo de rol y sesgo del profesor, propuestos por Dee (2007) y analizados también por Paredes.

Para este estudio se trabajará principalmente con datos de estudiantes de 8° básico, decisión respaldada por Valiente et al. (2019), quienes muestran que, durante los primeros grados de la educación primaria, la calidad de la relación entre estudiantes y docentes influye significativamente en el rendimiento académico: en lectura para los niños y en matemáticas para las niñas. Elegir 8° básico permite trabajar con la edad más temprana en la que los estudiantes tienen docentes de distinto sexo, ya que en cursos más tempranos, como 4.° básico, un mismo docente suele impartir varias asignaturas, lo que reduce considerablemente la variación en el género docente entre materias para un mismo estudiante. Sin perjuicio de lo anterior, también se incorporarán los puntajes obtenidos por esta misma cohorte en 4° básico, los cuales permitirán controlar distintos sesgos que se explican con mayor detalle en el marco teórico.

En este sentido, la presente investigación no solo busca aportar al debate académico, sino también incentivar el desarrollo de estrategias pedagógicas orientadas a fortalecer la equidad de género y a mejorar la calidad educativa en Chile.

Para ello, se utilizarán las bases de datos proporcionadas por el SIMCE (Sistema de Medición de la Calidad de la Educación), una evaluación estandarizada aplicada en Chile a estudiantes de distintos niveles escolares. Estas bases contienen información sobre el rendimiento académico en distintas asignaturas, junto con antecedentes de los estudiantes, docentes, familias y establecimientos, lo que permite analizar la relación entre el género del profesor y el desempeño académico.

En este contexto, la econometría cobra relevancia como una herramienta que permite estudiar de manera empírica una problemática que, a simple vista, podría quedar limitada a apreciaciones subjetivas sobre las diferencias de género, ya que, mediante distintas especificaciones econométricas, es posible cuantificar la relación entre el género del docente y el rendimiento académico, considerando al mismo tiempo otros factores que también pueden influir en el desempeño de los estudiantes y fortaleciendo, de esta forma, la evidencia obtenida a partir de los datos (Gujarati et al., 2004).

En línea con lo anterior, el presente estudio busca aportar evidencia empírica sobre la relación entre el género del docente y el rendimiento académico mediante la estimación de modelos econométricos, manteniendo una perspectiva de género que permita analizar si esta relación se manifiesta de manera diferente entre estudiantes hombres y mujeres, así como explorar los mecanismos que podrían estar detrás de estas diferencias.

Finalmente, para orientar y delimitar el proyecto se plantean las siguientes preguntas:

- ¿Existe un efecto causal del género del docente sobre el rendimiento académico de los estudiantes?
- ¿El efecto del género docente sobre el rendimiento académico difiere según la asignatura y el género del estudiante?
- ¿Cuál de los mecanismos analizados, modelo de rol o sesgo docente, resulta más consistente con el efecto observado del género docente sobre el rendimiento académico?
- ¿Qué implicancias pueden tener las diferencias encontradas en el efecto del género docente sobre estudiantes hombres y mujeres, particularmente respecto del papel de los referentes femeninos y de las desigualdades de género presentes en el ámbito educativo?

3 Objetivos

3.1 Objetivo General

Analizar el impacto del emparejamiento por género entre docente y estudiante en el rendimiento académico en Chile, con el fin de identificar cuál de los mecanismos analizados, modelo de rol o sesgo docente, resulta más consistente con el efecto observado.

3.2 Objetivos Específicos

1. Identificar características relevantes de los estudiantes, sus familias y docentes para el análisis del efecto del género docente sobre el rendimiento académico, construyendo a partir de ellas una base de datos limpia, integrada y consistente.
2. Estimar el efecto del género del docente sobre el rendimiento académico en Matemática y Ciencias Naturales, analizando si este difiere según el género del estudiante y la asignatura evaluada.
3. Desarrollar y contrastar distintas especificaciones econométricas para evaluar la robustez del efecto estimado del género docente, incorporando progresivamente variables de control y estrategias de efectos fijos que permitan reducir posibles sesgos asociados a heterogeneidad no observada.
4. Identificar cuál de los mecanismos analizados, modelo de rol o sesgo docente, resulta más consistente con el efecto observado del género docente sobre el rendimiento académico, mediante distintas estrategias empíricas orientadas a evaluar sus principales predicciones.
5. Analizar las implicancias de los resultados desde una perspectiva de género, particularmente respecto del papel de los referentes femeninos y de las desigualdades de género presentes en el



ámbito educativo.

4 Marco Teórico

4.1 Dinámicas de género en el aula

La presente investigación incorpora una perspectiva de género, entendida como una herramienta que permite comprender de qué manera las experiencias y oportunidades de hombres y mujeres se encuentran condicionadas por normas, estereotipos y relaciones de poder construidas socialmente. En este sentido, más que centrarse exclusivamente en las diferencias biológicas entre los sexos, este enfoque permite analizar cómo dichas construcciones pueden generar desigualdades en distintos ámbitos de la vida, entre ellos, las trayectorias educativas (Miralles-Cardona et al., 2020).

Dentro de este proceso, la escuela desempeña un papel relevante en la formación de identidades y en la manera en que el estudiantado comprende los roles asociados a hombres y mujeres, ya que el sistema educativo puede tanto contribuir a cuestionar estas diferencias como a reproducirlas mediante el currículo oculto; es decir, a través de normas, expectativas y prácticas que se transmiten de manera implícita en la vida escolar. En este contexto, el profesorado ocupa una posición central, puesto que no solo transmite contenidos académicos, sino que también participa en la construcción social del género dentro del aula (Rodríguez-Olay, 2023).

Por lo tanto, el interés de este estudio se centra en analizar de qué manera la interacción entre docentes y estudiantes influye en el rendimiento académico, particularmente en cómo el género del profesorado se relaciona con el desempeño del estudiantado en asignaturas donde los estereotipos de género han tenido históricamente mayor presencia, como Matemáticas y Ciencias.

Esta discusión parte de la idea de que la dinámica del aula no es neutra, por lo que distintos autores han estudiado la interacción entre el género de docentes y estudiantes. Entre ellos, Dee (2007) encontró que la coincidencia de género se relacionaba con mejores resultados académicos y con per-

cepciones más favorables dentro del aula. Sin embargo, la evidencia posterior no ha sido concluyente, pues Cho (2012), mediante modelos con efectos fijos por estudiante aplicados a países de la OCDE, encuentra escaso respaldo para la hipótesis de que el alumnado se beneficie sistemáticamente al tener docentes de su mismo género, mientras que Lim y Meer (2017), a partir de la asignación aleatoria de estudiantes a cursos de educación secundaria en Corea del Sur, muestran que las alumnas obtienen mejores resultados cuando son asignadas a profesoras y que los estudiantes hombres presentan pocos cambios.

En Chile, el principal antecedente corresponde al estudio de Paredes (2014), quien, utilizando datos de estudiantes de 8.º básico, encuentra que las alumnas se benefician de ser asignadas a profesoras, mientras que no observa efectos estadísticamente significativos sobre los alumnos. Además, sus resultados sugieren que este efecto no se explicaría principalmente por un trato diferenciado de las docentes hacia las estudiantes, sino por la influencia que una profesora puede ejercer como referente femenino. A partir de la discusión desarrollada por Dee, la autora propone un marco analítico para distinguir entre dos posibles mecanismos detrás de esta relación: el modelo de rol y el sesgo docente.

4.1.1. Mecanismos subyacentes: modelo de rol y sesgo docente

Paredes (2014) plantea que el género del profesor puede influir en el rendimiento debido a la reacción del estudiante ante el docente como referente (modelo de rol), o bien por un trato diferenciado del docente hacia el alumnado (sesgo). Este planteamiento permite establecer qué resultados serían consistentes con cada mecanismo y, a partir de ello, orientar el análisis empírico.

Para representar estos mecanismos de manera analítica y derivar predicciones que puedan ser contrastadas empíricamente, Paredes (2014) recurre a una formalización económica simplificada de las interacciones entre docentes y estudiantes.

En el primer mecanismo, el sesgo docente, se asume que el profesor distribuye su tiempo total de enseñanza de manera no neutral, priorizando al grupo de su mismo género. Por ello, este mecanismo puede explorarse observando si las expectativas de profesoras y profesores varían de forma distinta según la composición de género del curso y si el efecto asociado al género docente disminuye en aulas segregadas por género.

Para representarlo teóricamente, se introduce una función de utilidad docente:

$$V_t = \sum_{i=1}^2 a_{it} \frac{N_i}{N_1 + N_2} U(h_{it}) \quad \text{s.a.} \quad \bar{h} = h_{1t} + h_{2t} \quad (1)$$

donde h_{it} es el tiempo que el docente de género t dedica al grupo de estudiantes de género i , N_i es el número de estudiantes de ese grupo y a_{it} recoge la valoración que el docente asigna a dicho grupo. Si existe sesgo, el parámetro a_{it} reflejará una preferencia relativa por interactuar más intensamente con estudiantes de su propio género.

En el segundo mecanismo, el modelo de rol, el rendimiento depende también de la identificación simbólica del estudiante con el docente. Por ello, este efecto está condicionado por el entorno familiar y social, siendo más fuerte cuando el estudiante carece de referentes sólidos en su hogar o se enfrenta a asignaturas con expectativas sociales desfavorables para su género.

Analíticamente, se expresa como:

$$S_{it} = l(f_i(h_{it}, h_{jt}), b_i r_{iths}, ability_i) + \varepsilon_i \quad (2)$$

donde S_{it} es el desempeño del estudiante, $l(\cdot)$ representa la función mediante la cual los distintos factores se traducen en dicho desempeño y $f_i(\cdot)$ describe la manera en que el tiempo de enseñanza se transforma en aprendizaje para el grupo de estudiantes de género i , incorporando tanto el tiempo

dedicado directamente a dicho grupo, h_{it} , como el posible aprendizaje indirecto que puede surgir del tiempo asignado al otro grupo de estudiantes, de género j ; sin embargo, se asume que el efecto del tiempo dedicado directamente es mayor. Además, b_i representa el grado en que el desempeño del grupo i responde al efecto de modelo de rol, $ability_i$ representa la habilidad previa del estudiante y r_{iths} captura dicho efecto, el cual depende del docente (t), del hogar (h) y del contexto social (s).

La importancia de esta distinción radica en que ambos mecanismos conducen al mismo resultado observable (por ejemplo, una mejora en el rendimiento de las niñas con profesoras), pero implican interpretaciones muy diferentes sobre lo que ocurre en el sistema educativo.

4.2 Contexto educativo chileno

En línea con lo planteado en la sección anterior, también es importante considerar el contexto educativo en el que se desarrolla la relación entre el género docente y el rendimiento estudiantil, pues las normas y formas de organización del sistema configuran las condiciones en que esta ocurre. Teniendo esto en cuenta, cabe señalar que, a lo largo de su historia, el sistema educativo chileno ha experimentado profundas modificaciones, tanto en materia de gobernanza como en sus mecanismos de financiamiento y evaluación. En particular, dado que la evidencia utilizada como punto de comparación se basa en datos SIMCE de 2009 (Paredes, 2014), esta sección busca contextualizar el análisis de la presente investigación a partir de las transformaciones experimentadas desde ese año hasta el período analizado.

Entre las transformaciones ocurridas durante este período, uno de los principales hitos fue la promulgación de la Ley General de Educación (LGE) en 2009 (Ministerio de Educación de Chile, 2009), la cual reemplazó a la Ley Orgánica Constitucional de Enseñanza (LOCE) y estableció un nuevo marco normativo para la educación escolar, ya que incorporó principios fundamentales como

calidad, equidad, inclusión y participación, asignando al Estado un rol más activo como garante del derecho a una educación integral. De este modo, redefinió el marco institucional en que tiene lugar la interacción entre docentes y estudiantes, al desplazar parcialmente la lógica puramente competitiva que había caracterizado al sistema en décadas anteriores y reforzar una concepción de la educación más vinculada al derecho social.

Por otro lado, la implementación de la Ley de Inclusión Escolar en 2015 (Ministerio de Educación de Chile, 2015), que prohibió el lucro, la selección académica y el copago en establecimientos con financiamiento público, también cobra relevancia porque incidió sobre la composición del estudiantado y sobre la distribución de los alumnos entre establecimientos, puesto que esta normativa buscó desincentivar la segmentación escolar y promover un acceso más equitativo a las oportunidades educativas, especialmente para estudiantes de contextos vulnerables. De forma complementaria, a partir de 2016 comenzó a operar el Sistema de Admisión Escolar (SAE) (Ministerio de Educación de Chile, 2018), el cual introdujo un algoritmo centralizado de asignación que reemplazó los mecanismos de selección discrecional utilizados hasta entonces, con el objetivo de asegurar procesos más transparentes y justos.

El SAE, en específico, resulta importante para este estudio porque reemplazó un proceso descentralizado, en el que los establecimientos podían seleccionar estudiantes y las familias debían postular de manera individualizada, por un mecanismo centralizado de asignación basado en aceptación diferida. Con ello, se redujo la posibilidad efectiva de selección por parte de los establecimientos y se atenuó la discrecionalidad en la admisión, mientras que la posibilidad de postular simultáneamente a varios establecimientos y la revelación de las verdaderas preferencias como estrategia dominante disminuyeron los incentivos a adoptar comportamientos estratégicos durante el proceso de postulación (Ministerio de Educación de Chile, 2018). Ahora bien, cabe destacar que esto no implica una asignación completamente aleatoria de estudiantes ni, menos aún, de profesores, por lo que no permi-

te asumir una ausencia total de selección. Sin embargo, su implementación resulta relevante para la presente investigación, ya que el análisis se desarrolla en un sistema de admisión más transparente y estandarizado que el existente en 2009.

Sin perjuicio de lo anterior, resulta relevante reconocer que, a pesar de estos cambios, el sistema escolar chileno aún opera en un contexto de persistentes desigualdades de género, lo cual se evidencia tanto en la baja participación femenina en carreras STEM como en la propia estructura ocupacional del sistema educativo. En efecto, si bien las mujeres constituyen el 78,1 % del profesorado en educación básica y este porcentaje asciende al 99,1 % en los primeros niveles del ciclo escolar, su participación en cargos de liderazgo sigue siendo limitada, pues solo el 35,6 % de los cargos directivos están ocupados por mujeres (Centro de Estudios, 2024). Esta disparidad refleja una segmentación vertical en la estructura profesional del sistema educativo, donde las mujeres presentan una amplia participación en la docencia, pero continúan subrepresentadas en posiciones de liderazgo y toma de decisiones.

Finalmente, en este contexto, resulta especialmente relevante contar con instrumentos que permitan observar el desempeño escolar de manera sistemática y comparable. En este sentido, herramientas como el SIMCE adquieren relevancia no solo para monitorear la evolución del desempeño estudiantil, sino también para identificar brechas persistentes. Precisamente por ello, en la siguiente sección se profundiza en el concepto de rendimiento académico y en el uso del SIMCE como medida central para esta investigación.

4.3 Evaluación del rendimiento académico a través del SIMCE

El SIMCE es una evaluación que contempla pruebas en Lenguaje, Matemática, Ciencias e Historia, aplicadas, según el calendario de evaluación correspondiente, en 4º, 6º y 8º básico y 2º medio, a las que se suman cuestionarios contextuales dirigidos a estudiantes, apoderados, docentes y equipos

directivos, los cuales permiten recoger información sobre el entorno escolar, las prácticas pedagógicas y las condiciones socioeconómicas del alumnado. De esta forma, constituye una herramienta estandarizada que permite evaluar los logros de aprendizaje del estudiantado en distintos momentos de su trayectoria escolar y en distintas áreas relevantes del currículo nacional y, al mismo tiempo, vincular dichos resultados con información contextual sobre estudiantes, apoderados, docentes y establecimientos (Agencia de Calidad de la Educación, s.f.). En base a lo anterior, resulta una fuente especialmente útil para la presente investigación, ya que no solo entrega puntajes en asignaturas como Matemática y Ciencias, sino que también permite relacionarlos con características del entorno escolar y familiar.

En este marco, el rendimiento académico se entiende como el nivel de logro que alcanzan los estudiantes en relación con los objetivos de aprendizaje establecidos en el currículo escolar y suele medirse mediante calificaciones, pruebas estandarizadas, evaluaciones formales y otros instrumentos diseñados para cuantificar los conocimientos, habilidades y competencias adquiridas en distintas asignaturas (Gustafsson et al., 2018).

Ampliando esta definición, el rendimiento académico no solo refleja la capacidad cognitiva del estudiante, sino que también está influido por factores individuales, familiares, escolares y contextuales, entre los cuales destacan la calidad de la enseñanza, la motivación, el entorno socioeconómico, el capital cultural del hogar, las condiciones del establecimiento educativo y el vínculo con el profesorado (Coleman, 1968).

Sin embargo, cabe reconocer que, si bien el SIMCE constituye una medida estandarizada y contextualizada del rendimiento académico, sus resultados deben interpretarse considerando que las diferencias observadas entre hombres y mujeres pueden estar asociadas a factores que trascienden únicamente los conocimientos adquiridos. Esto puede observarse, por ejemplo, en el caso chileno, donde, a partir de datos de PISA (Programme for International Student Assessment) 2012 y de los Cuestiona-

rios de Calidad y Contexto Educativo del SIMCE 2013, la Agencia de Calidad de la Educación reporta que, en general, los hombres se declaran más motivados para aprender Matemática, más seguros de sus habilidades, con menor estrés y con mayor percepción de valoración por parte de sus profesores (Agencia de Calidad de la Educación, 2016).

Esto resulta consistente con la literatura sobre amenaza del estereotipo, la cual ha mostrado que el desempeño en pruebas estandarizadas de Matemáticas puede verse afectado cuando las estudiantes rinden bajo condiciones en las que se activa el estereotipo de que las mujeres tendrían un peor desempeño en esta área. Este efecto puede observarse, por ejemplo, en el estudio de Spencer et al. (1999), donde mujeres con alta capacidad matemática obtuvieron resultados inferiores a los de los hombres cuando la prueba fue presentada como una evaluación que históricamente mostraba diferencias por género; en cambio, esa brecha desapareció cuando la prueba fue descrita como una medición sin diferencias entre hombres y mujeres.

En consecuencia, si bien el SIMCE permite medir logros de aprendizaje de manera comparable, en una investigación como esta resulta importante considerar que los puntajes también pueden verse influidos por factores emocionales y simbólicos que no afectan por igual a estudiantes de distinto género, especialmente en asignaturas donde persisten estereotipos sobre capacidad académica.

4.4 Modelo econométrico

A partir de los antecedentes desarrollados previamente, el análisis econométrico constituye la herramienta mediante la cual se busca estimar el efecto del género docente sobre el rendimiento académico e integrar la perspectiva de género dentro del análisis empírico, al expresar la relación entre género docente, género del estudiante y desempeño académico mediante parámetros que pueden ser estimados y contrastados a partir de los datos.

En esta línea, el análisis toma como referencia estudios como Dee (2007) y Paredes (2014), que analizan los efectos asociados al emparejamiento por género entre docentes y estudiantes, por ende la estrategia comienza con un modelo de regresión multivariable que incorpora progresivamente variables relevantes para el desempeño académico, manteniendo como parámetro central la interacción entre el género del estudiante y el del docente.

Un modelo de regresión multivariable es una herramienta que permite analizar la relación entre una variable de interés y un resultado, incorporando simultáneamente otros factores que también pueden estar asociados con dicho resultado; de esta manera, permite estimar la relación de interés manteniendo constantes, en la medida en que son observadas e incorporadas al modelo, otras características potencialmente relevantes, conocidas como variables de control (Wooldridge, 2019; Angrist y Pischke, 2008).

Aplicada al presente estudio, esta especificación puede expresarse como:

$$Y_{is}^z = \beta_0 + \beta_1 \text{Mujer}_i + \beta_2 \text{Profesora}_{is} + \beta_3 (\text{Mujer}_i \times \text{Profesora}_{is}) + \rho \text{PuntajePrevio}_{is}^z + \delta \text{Promedio}_i^z + \gamma' F_i + \theta' T_{is} + \varepsilon_{is}, \quad (3)$$

donde Y_{is}^z representa el puntaje estandarizado obtenido por el estudiante i en la asignatura s ; Mujer_i identifica a las estudiantes mujeres y Profesora_{is} indica que el docente de la asignatura es mujer. Por su parte, $\text{PuntajePrevio}_{is}^z$ y Promedio_i^z permiten considerar el desempeño académico previo, mientras que F_i reúne características familiares relevantes y T_{is} incorpora características del docente que pueden variar entre asignaturas. Dentro de esta formulación, β_2 representa el efecto asociado a ser asignado a una profesora para los estudiantes hombres, utilizados como categoría de referencia, mientras que β_3 permite determinar cuánto difiere dicho efecto para las estudiantes mujeres.

La incorporación progresiva de estas variables busca disminuir el riesgo de atribuir al género

docente diferencias asociadas a características observables de los estudiantes y de su entorno; sin embargo, aun después de incorporar controles, pueden persistir factores no observados relacionados tanto con el rendimiento como con la asignación del docente, por lo que resulta necesario avanzar hacia una estrategia que permita controlar parte de esta heterogeneidad no observada.

4.4.1. Ajuste mediante efectos fijos

Los modelos de efectos fijos permiten controlar aquellas características, observadas o no observadas, que permanecen constantes dentro de una determinada unidad de análisis, utilizando para la estimación la variación que ocurre al interior de dicha unidad (Wooldridge, 2019; Angrist y Pischke, 2008). En términos simples, cada unidad funciona como su propio punto de comparación, de modo que, en lugar de atribuir las diferencias existentes entre unidades únicamente a las variables incluidas en el modelo, esta estrategia concentra el análisis en los cambios que ocurren dentro de una misma unidad, manteniendo constantes aquellas características propias que no varían entre las observaciones.

En el caso de esta investigación, la estrategia principal incorpora efectos fijos por estudiante, aprovechando que un mismo alumno es observado tanto en Matemática como en Ciencias Naturales y puede encontrarse expuesto a docentes de distinto género entre ambas asignaturas. De esta forma, la comparación deja de realizarse principalmente entre estudiantes distintos y pasa a utilizar las diferencias observadas dentro de un mismo estudiante, por lo que características individuales que permanecen constantes entre ambas asignaturas, sean observadas o no, como las condiciones familiares persistentes o ciertos rasgos individuales, quedan absorbidas por el efecto fijo individual, reduciendo así la posibilidad de que diferencias permanentes entre alumnos sean confundidas con el efecto asociado al género docente.

Adicionalmente, con el objetivo de evitar que diferencias asociadas a estos factores sean atri-

buidas al género docente, concentrando la comparación en estudiantes expuestos a un mismo contexto de enseñanza, se incorpora una especificación con efectos fijos por curso-asignatura, que permite controlar características comunes a los estudiantes que comparten un mismo establecimiento, curso y asignatura, tales como aspectos propios del grupo y características del docente que permanecen constantes dentro de dicha unidad, como su experiencia o formación.

Bajo esta estrategia, la especificación general puede expresarse como:

$$\begin{aligned} Y_{is}^z = & \alpha_i + \mu_s + \beta_1 Profesora_{is} + \beta_2 (Mujer_i \times Profesora_{is}) \\ & + \rho PuntajePrevio_{is}^z + \phi (Mujer_i \times PuntajePrevio_{is}^z) + \theta' T_{is} \\ & + \lambda (Mujer_i \times Ciencias_s) + \varepsilon_{is}, \end{aligned} \quad (4)$$

donde α_i corresponde al efecto fijo del estudiante y absorbe aquellas características observadas y no observadas que permanecen constantes entre asignaturas, mientras que μ_s controla diferencias sistemáticas entre Matemática y Ciencias Naturales. El vector T_{is} reúne características docentes que pueden variar entre ambas observaciones, mientras que la interacción entre el género del estudiante y su rendimiento previo permite considerar posibles diferencias en la relación entre desempeño anterior y actual según género. Asimismo, la interacción entre el género del estudiante y Ciencias Naturales permite que la diferencia de rendimiento entre hombres y mujeres pueda variar entre Matemática y Ciencias Naturales.

Como punto de comparación para esta estrategia, la investigación considera además una especificación base con efectos fijos por establecimiento, la cual concentra la comparación entre estudiantes pertenecientes a un mismo colegio, de modo que las características comunes y relativamente constantes del establecimiento, como aspectos de su organización, infraestructura o contexto institucional, quedan controladas. Así, las diferencias permanentes entre establecimientos no se utilizan para identificar el

efecto del género docente; no obstante, esta especificación no controla las diferencias no observadas que pueden existir entre estudiantes de un mismo colegio, por lo que funciona principalmente como una base de comparación antes de avanzar hacia la estimación con efectos fijos por estudiante.

Bajo estas formulaciones, el parámetro de mayor interés continúa siendo la interacción entre el género del estudiante y el género docente, ya que permite establecer si la exposición a una profesora presenta un efecto diferencial sobre el rendimiento de las estudiantes mujeres respecto de sus compañeros hombres.

Una vez estimado este efecto, su interpretación requiere considerar que las diferencias observadas no se producen de manera aislada, sino dentro de un contexto educativo atravesado por expectativas, estereotipos y relaciones de género, aspecto que se desarrolla en la siguiente sección.

4.5 Análisis e interpretación de los resultados con perspectiva de género

A partir de la estrategia econométrica anterior, la perspectiva de género orienta la forma en que serán interpretadas las diferencias estimadas entre estudiantes hombres y mujeres. En este sentido, tal como plantea Scott (1986), el género no se incorpora únicamente como una característica utilizada para distinguir grupos dentro de la muestra, sino como una dimensión que estructura relaciones sociales, expectativas y oportunidades dentro del espacio educativo; por ello, una diferencia en el efecto del género docente no será interpretada como el reflejo de capacidades inherentes a hombres y mujeres, sino considerando las condiciones sociales y escolares en las que dicha diferencia se produce.

Esta consideración resulta especialmente relevante al interpretar el mecanismo de modelo de rol, ya que la identificación de una estudiante con una profesora no ocurre de manera aislada, sino dentro de un contexto en el que los estereotipos de género y la representación desigual de hombres y mujeres pueden influir en las expectativas que los estudiantes construyen respecto de sus propias

capacidades y posibilidades de éxito. Bajo esta perspectiva, Morgenroth et al. (2015) plantean que los modelos de rol pueden influir sobre la motivación y las metas de las personas al actuar como ejemplos de comportamiento, representar que determinados logros son posibles y constituir una fuente de inspiración; bajo esta lógica, la presencia de una profesora puede adquirir una relevancia particular para las estudiantes cuando existen menos referentes femeninos en su entorno o cuando se encuentran en espacios tradicionalmente asociados a los hombres, no simplemente por compartir el mismo género, sino porque su presencia puede contribuir a ampliar aquello que las estudiantes perciben como posible para sí mismas.

A partir de esta interpretación, si el efecto del género docente responde a un mecanismo de modelo de rol, cabría esperar que su magnitud varíe de acuerdo con las condiciones en que esa identificación puede resultar más o menos relevante. Por esta razón, el análisis considera las diferencias entre Matemática y Ciencias Naturales, la educación de la madre y dimensiones asociadas a la percepción de las propias capacidades, en particular, una mayor relevancia del efecto en contextos donde existen menos referentes femeninos con trayectoria educativa, o una asociación positiva entre la presencia de una profesora y la percepción de capacidad de las estudiantes, constituirían patrones consistentes con este mecanismo. Esta lógica se encuentra también presente en Paredes (2014), quien plantea que el efecto de una profesora debería ser mayor en asignaturas tradicionalmente masculinizadas y entre niñas que cuentan con menos modelos femeninos fuertes fuera de la escuela.

Por otra parte, la interpretación del mecanismo de sesgo docente parte de reconocer que los propios profesores se encuentran insertos en el mismo contexto social y, por tanto, sus expectativas y formas de interacción también pueden verse influidas por estereotipos acerca de las capacidades de hombres y mujeres. Estos sesgos no requieren necesariamente un trato discriminatorio deliberado, sino que pueden manifestarse a través de diferencias en las expectativas, la retroalimentación, las re-

comendaciones o las interacciones que se establecen con los estudiantes. En esta línea, Carlana (2019) encuentra que la exposición a docentes de Matemática con estereotipos de género más fuertes amplía la brecha de rendimiento entre niños y niñas y se relaciona con una menor confianza de las estudiantes en sus propias capacidades matemáticas, mostrando que las creencias del docente pueden trasladarse a las experiencias y percepciones académicas del alumnado.

Bajo esta lógica, el análisis del sesgo docente considera principalmente las expectativas de los profesores y la composición de género del aula, ya que ambas dimensiones permiten observar manifestaciones que serían compatibles con un tratamiento diferencial entre estudiantes hombres y mujeres. Si profesores y profesoras modifican de manera distinta sus expectativas frente a la composición de género del curso, ello podría reflejar diferencias en las valoraciones que realizan sobre las capacidades de sus estudiantes; del mismo modo, si el efecto del género docente se origina en un trato diferencial entre ambos grupos dentro de la sala de clases, cabría esperar que este adquiriera mayor relevancia en aulas mixtas, donde estudiantes hombres y mujeres se encuentran expuestos simultáneamente al mismo docente (Paredes, 2014).

De esta manera, la estrategia econométrica y la perspectiva de género se articulan para estimar e interpretar las diferencias asociadas al género docente, permitiendo evaluar si los patrones observados resultan más consistentes con un mecanismo de modelo de rol o de sesgo docente. Ahora bien, es importante considerar que este conjunto de análisis no permite demostrar de manera aislada la existencia de uno u otro mecanismo, pero sí contrastar sus principales predicciones a partir de la evidencia empírica.

5 Consolidación de datos

Antes de desarrollar la estrategia econométrica, fue necesario construir una base de datos analítica que integrara la información disponible para los estudiantes, sus familias y docentes. Para ello, se utilizó principalmente información proveniente del SIMCE 2017, correspondiente a estudiantes de 8° básico, la cual fue complementada con antecedentes académicos previos de la misma cohorte obtenidos a partir del SIMCE 2013 y de las calificaciones escolares de 2016 disponibles en el Centro de Estudios del Ministerio de Educación.

La selección de los años y asignaturas utilizadas responde tanto al objetivo del estudio como a la disponibilidad de información necesaria para realizar el seguimiento de una misma cohorte. En particular, las bases disponibles permiten identificar a estudiantes que rindieron el SIMCE de 4° básico en 2013 y que posteriormente fueron evaluados en 8° básico durante 2017, concentrándose el análisis en Matemática y Ciencias Naturales, asignaturas para las cuales es posible contar con información comparable entre ambos períodos. De esta forma, el año 2017 constituye el período principal de análisis, mientras que los puntajes obtenidos por la misma cohorte en 2013 permiten incorporar antecedentes previos específicos para cada asignatura.

De manera complementaria, se incorporaron las calificaciones escolares correspondientes a 2016 con el propósito de contar con una medida más cercana temporalmente del desempeño académico general de los estudiantes. Así, mientras los puntajes SIMCE de 2013 permiten aproximar diferencias previas en Matemática y Ciencias Naturales, las calificaciones de 2016 aportan información adicional sobre la trayectoria académica inmediatamente anterior al período analizado.

A partir de estas fuentes, la construcción de la base final requirió procesos de depuración, homologación y vinculación de la información, orientados a asegurar la consistencia de los registros y su posterior utilización en el análisis econométrico. En las secciones siguientes se describe en detalle

cada una de estas etapas.

5.1 Especificación de bases de datos utilizadas del SIMCE 2017

Dado que el SIMCE 2017 constituye la principal fuente de información para el análisis, su consolidación requirió integrar distintas bases que contienen antecedentes de estudiantes, sus familias y docentes. En esta sección se describen las fuentes utilizadas y los principales criterios aplicados para su depuración, con el propósito de conformar posteriormente una base única y consistente para el análisis.

5.1.1. Base de alumnos

La base de alumnos contiene información individual de los estudiantes que rindieron la evaluación SIMCE 2017, de la cual se seleccionaron las variables presentadas en la Tabla 1.

Tabla 1: Variables seleccionadas desde la base de alumnos

Variable	Descripción	Tipo
agno	Año de la evaluación	N Numérica
mrún	Identificador único MINEDUC para uso longitudinal	N Numérica
idalumno	Identificador único SIMCE para uso transversal	N Numérica
gen_alu	Género del estudiante	C Categórica
rbd	Rol Base de Datos del establecimiento	N Numérica
grado	Grado que rinde la evaluación SIMCE	C Categórica
cod_curso	Código de curso para vinculación con cuestionarios docentes	N Numérica
ptje_mate8b_alu	Puntaje del estudiante en SIMCE Matemática	N Numérica
ptje_nat8b_alu	Puntaje del estudiante en SIMCE Ciencias Naturales	N Numérica

La limpieza de esta base se orientó a conservar únicamente los registros con información válida en el identificador longitudinal y en los puntajes de Matemática y Ciencias Naturales, tras lo cual la muestra pasó de 247.495 a 199.832 observaciones, lo que representa una reducción de 47.663 registros, equivalente al 19,3 % de la base inicial.

5.1.2. Cuestionarios docentes

Para incorporar información sobre las características de los docentes, se utilizaron de manera separada los cuestionarios correspondientes a Matemática y Ciencias Naturales, asignaturas seleccionadas por la disponibilidad de archivos específicos con información del profesor responsable a nivel de curso. A partir de estas bases se seleccionaron las variables presentadas en la Tabla 2.

Tabla 2: Variables seleccionadas y construidas a partir de los cuestionarios docentes

Variable	Descripción y codificación	Tipo
agno	Año correspondiente a la evaluación.	Numérica
rbd	Rol Base de Datos que identifica al establecimiento educacional.	Numérica
grado	Grado correspondiente a la evaluación SIMCE.	Categórica
cod_curso	Código que identifica al curso dentro del establecimiento y permite vincular la información del docente con los estudiantes.	Numérica
genero_profnat/mate	Género del docente: 1 = profesor; 2 = profesora.	Categórica
edad_profnat/mate	Edad del docente, expresada en años.	Numérica
titulo_profnat/mate	Indicador de titulación docente: 0 = no titulado; 1 = titulado.	Dicotómica
especializacion_profnat/mate	Indicador de especialización en la asignatura correspondiente: 0 = no especializado; 1 = especializado.	Dicotómica
experiencia_prof_nat/mate	Años de experiencia docente.	Numérica
expect_profnat/mate	Expectativa educacional del docente respecto del nivel educativo que espera que alcance la mayoría de sus estudiantes, medida en una escala de 1 a 6, donde valores mayores representan expectativas más altas.	Ordinal
INDCOND_NAT/MATE	Índice estandarizado de condiciones laborales.	Continua
INMIND_NAT/MATE	Índice estandarizado de mentalidad docente.	Continua

Ambos cuestionarios docentes fueron sometidos a un procedimiento de depuración equivalente, con el propósito de aplicar criterios comparables a las bases de Matemática y Ciencias Naturales antes de su incorporación a la base final. Como criterio general, las respuestas faltantes o codificadas como inválidas fueron excluidas cuando impedían identificar o construir correctamente las variables

de interés. En particular, dependiendo de la pregunta, códigos como 0 y 99 correspondían a ausencia de respuesta o registros no válidos. Adicionalmente, para algunas variables fue necesario establecer criterios específicos de consistencia y recodificación, los cuales se describen a continuación:

- **Edad del docente:** se conservaron únicamente los registros correspondientes a docentes con edades entre 22 y 80 años, con el propósito de excluir valores fuera de un rango plausible.
- **Género del docente:** se mantuvieron únicamente las observaciones correspondientes a las dos categorías válidas del cuestionario, donde 1 identifica a profesores hombres y 2 a profesoras mujeres.
- **Titulación docente:** a partir de las preguntas asociadas a la formación profesional se construyó una variable dicotómica que distingue entre docentes titulados y no titulados, considerando como inconsistentes aquellos registros que indicaban simultáneamente ambas condiciones o que no permitían determinar ninguna de ellas.
- **Especialización disciplinar:** se identificó si el docente contaba con especialización en la asignatura correspondiente, diferenciando entre quienes declaraban formación específica en Matemática o Ciencias Naturales y quienes presentaban especialización en otras áreas.
- **Experiencia docente:** la información disponible distingue entre docentes con menos de un año de experiencia y aquellos que reportan directamente el número de años ejercidos. Para construir una medida continua, los primeros fueron codificados con valor 0, mientras que para los demás se conservaron únicamente valores positivos inferiores a 60 años.
- **Expectativas docentes:** se conservaron las seis categorías originales de respuesta, donde valores mayores representan expectativas educacionales más altas.

Por otro lado, respecto a los índices de condiciones laborales y mentalidad docente, dado que su elaboración requirió procedimientos específicos de depuración, recodificación y estandarización, su construcción se presenta en la Sección 5.2.

La tabla 3 resume la cantidad de registros descartados sucesivamente durante este proceso y el número de observaciones disponible al finalizar la depuración de cada cuestionario.

Tabla 3: Depuración de los cuestionarios docentes

Etapa	Matemática	Ciencias Naturales
Observaciones iniciales	8.324	8.186
Edad faltante o fuera del rango establecido	99	100
Género inválido	4	11
Titulación sin información válida	23	23
Especialización sin información válida	469	460
Experiencia sin información o inconsistente	774	825
Expectativa docente sin información válida	176	222
Información insuficiente para condiciones laborales	122	181
Información insuficiente para mentalidad docente	153	189
Total de observaciones eliminadas	1.820	2.011
Observaciones finales	6.504	6.175

Finalmente, las bases de docentes de Matemática y Ciencias Naturales pasaron, respectivamente, de 8.324 y 8.186 observaciones a 6.504 y 6.175 registros, lo que equivale a reducciones del 21,9 % y 24,6 % de sus muestras iniciales.

5.1.3. Cuestionario de estudiantes

Para incorporar información sobre las creencias y percepciones de los estudiantes, se utilizaron los cuestionarios aplicados a los alumnos durante el SIMCE 2017, cuya información se encontraba distribuida en dos formularios, identificados como A y B. Dado que estos no contenían exactamente al mismo conjunto de estudiantes, fueron unidos verticalmente para conformar una única base de trabajo,

manteniendo el identificador de cada alumno para su posterior vinculación con las demás fuentes de información.

A partir de estos cuestionarios se seleccionaron las preguntas necesarias para construir tres variables utilizadas posteriormente en el análisis. Por una parte, se elaboraron medidas de mentalidad general y mentalidad hacia Matemática, orientadas a recoger las creencias de los estudiantes respecto de la posibilidad de desarrollar sus capacidades mediante el aprendizaje. Por otra, se construyó un indicador de autoconcepto matemático, destinado a representar la percepción que tiene el estudiante respecto de sus propias capacidades para aprender y desenvolverse en esta asignatura.

Como criterio general de depuración, las respuestas faltantes o codificadas como inválidas fueron tratadas como valores perdidos, mientras que las respuestas válidas fueron posteriormente recodificadas y combinadas de acuerdo con las características de cada indicador. Debido a que su construcción requirió criterios específicos relacionados con el sentido de las escalas, el número mínimo de respuestas disponibles y su posterior estandarización, el procedimiento correspondiente se presenta en la Sección 5.2.

Las variables provenientes del cuestionario de estudiantes utilizadas en el análisis se presentan en la Tabla 4.

Tabla 4: Variables provenientes del cuestionario de estudiantes utilizadas en el análisis

Variable	Descripción	Tipo
agno	Año correspondiente a la evaluación.	Numérica
grado	Grado correspondiente a la evaluación SIMCE.	Catórica
idalumno	Identificador del estudiante utilizado para vincular la información entre las distintas bases del SIMCE 2017.	Numérica
rbd	Rol Base de Datos que identifica al establecimiento educacional.	Numérica
cod_curso	Código que identifica al curso dentro del establecimiento.	Numérica
INMIND_GENERAL_CEST	Índice estandarizado de mentalidad general del estudiante.	Continua
INMIND_MATE_CEST	Variable estandarizada de mentalidad hacia Matemática.	Continua
autoconcepto_mate_z	Índice estandarizado de autoconcepto matemático.	Continua

A partir de la construcción de estos índices, la base consolidada redujo el número de observaciones disponibles desde 215.201 a 207.216 registros, lo que corresponde a una pérdida del 3,7 % de la muestra inicial.

5.1.4. Cuestionario de apoderados

Para incorporar información sobre el entorno familiar de los estudiantes, se utilizó el cuestionario de apoderados correspondiente al SIMCE 2017, a partir del cual se seleccionaron y construyeron las variables presentadas en la Tabla 5.

Tabla 5: Variables provenientes del cuestionario de apoderados utilizadas en el análisis

Variable	Descripción y codificación	Tipo
agno	Año correspondiente a la evaluación.	Numérica
idalumno	Identificador del estudiante utilizado para vincular la información con las demás bases del SIMCE 2017.	Numérica
rbd	Rol Base de Datos que identifica al establecimiento educacional.	Numérica
grado	Grado correspondiente a la evaluación SIMCE.	Categórica
cod_curso	Código que identifica al curso dentro del establecimiento.	Numérica
educ_padre	Disponibilidad de información sobre educación del padre: 0 = sin información válida; 1 = con información válida.	Dicotómica
educ_madre	Nivel educacional de la madre: 1 = básica o menos; 2 = media; 3 = superior.	Ordinal
ingreso_hogar_orig	Ingreso mensual del hogar, conservado según los 15 tramos originales del cuestionario.	Ordinal
ingreso_hogar_4	Ingreso mensual del hogar agrupado en cuatro niveles: 1 = bajo; 2 = medio; 3 = medio alto; 4 = alto.	Ordinal
expecta_apod	Expectativa educacional del apoderado: 1 = no completar educación media; 2 = completar educación media; 3 = educación superior.	Ordinal

Una vez seleccionadas las variables de interés, se procedió a revisar y depurar sus respuestas con el propósito de disponer de información consistente para el análisis. Como criterio general, los valores faltantes y las respuestas codificadas como 0 o 99 fueron consideradas inválidas y, por tanto,

tratadas como valores perdidos.

A diferencia de las demás variables, la educación del padre presentó una cantidad considerable de respuestas faltantes, por lo que eliminar estos registros habría reducido de manera importante la muestra disponible. Por esta razón, se construyó un indicador de disponibilidad de información, asignando el valor 1 cuando existía una respuesta válida sobre el nivel educacional del padre y 0 cuando esta información no se encontraba disponible, permitiendo conservar estas observaciones en la base.

La Tabla 6 detalla la cantidad de observaciones eliminadas.

Tabla 6: Depuración del cuestionario de apoderados

Motivo de eliminación	Observaciones
Educación de la madre inválida o faltante	7.285
Ingreso del hogar inválido o faltante	4.187
Expectativa educacional inválida o faltante	3.867
Total de observaciones eliminadas	15.339
Observaciones finales	176.588

Finalmente, la base de apoderados pasó de 191.927 a 176.588 observaciones, lo que representa una reducción de 15.339 registros, equivalente aproximadamente al 8,0 % de la muestra inicial.

5.2 Construcción de indicadores

Como se mencionó anteriormente, a partir de los cuestionarios de docentes y estudiantes se construyeron distintos indicadores que posteriormente fueron utilizados en el análisis. Si bien cada uno recoge dimensiones diferentes y presenta particularidades en su elaboración, algunos comparten procedimientos comunes, entre ellos la estandarización, aplicada cuando fue necesario expresar las variables construidas en una escala comparable:

$$Z_i = \frac{I_i - \bar{I}}{s_I}, \quad (5)$$

donde I_i corresponde al valor del indicador antes de su estandarización, $\bar{I}$ a su media muestral y s_I a su desviación estándar. Como resultado, la variable estandarizada presenta media cero y desviación estándar igual a uno, permitiendo interpretar sus valores en términos de desviaciones estándar respecto de la media de la muestra.

A partir de lo anterior, en esta sección se describen tanto los procedimientos comunes como las decisiones específicas adoptadas para la construcción de cada indicador.

5.2.1. Indicadores de mentalidad

Los indicadores de mentalidad docente y mentalidad general de los estudiantes fueron contruidos mediante un procedimiento equivalente haciendo uso de las preguntas presentadas en el Anexo 11.1. En ambos casos, las preguntas utilizadas presentan una escala de respuesta entre 1 y 4 y se encuentran formuladas de manera que una mayor puntuación refleja un mayor acuerdo con afirmaciones que consideran que la inteligencia y las capacidades son difíciles de modificar mediante el aprendizaje, por lo que las respuestas fueron invertidas mediante la Ecuación 6 para que valores más altos representaran una mayor creencia en que dichas capacidades pueden desarrollarse mediante el aprendizaje.

$$x_{ij}^* = 5 - x_{ij}, \quad (6)$$

donde x_{ij} corresponde a la respuesta original del individuo i al ítem j y x_{ij}^* representa su valor recodificado. De esta forma, una respuesta igual a 1 pasa a tomar el valor 4, mientras que una respuesta igual a 4 pasa a tomar el valor 1.

Una vez invertidas las respuestas, cada indicador fue obtenido mediante el promedio de los ítems válidos disponibles:

$$M_i = \frac{1}{m_i} \sum_{j \in V_i} x_{ij}^*, \quad (7)$$

donde V_i representa el conjunto de respuestas válidas disponibles para el individuo i y m_i corresponde al número de respuestas utilizadas en el cálculo.

Como criterio de construcción, para la mentalidad docente se requirieron al menos 3 de las 4 respuestas válidas, mientras que para la mentalidad general de los estudiantes se exigieron al menos 2 de las 3 respuestas utilizadas.

Finalmente, los promedios obtenidos fueron estandarizados mediante la Ecuación 5, de manera que valores más altos de ambos indicadores representan una mayor creencia en que la inteligencia y las capacidades pueden desarrollarse mediante el aprendizaje.

5.2.2. Indicadores asociados a Matemática

Además de los indicadores generales de mentalidad, se construyó una medida específica de mentalidad hacia Matemática a partir de la pregunta presentada en el Anexo 11.2. Esta medida busca recoger la creencia del estudiante respecto de la posibilidad de desarrollar su capacidad matemática mediante el aprendizaje y, a diferencia de los indicadores anteriores, se construyó a partir de una única pregunta.

Dado que la afirmación estaba formulada de manera que una mayor puntuación reflejaba una mayor creencia en que la capacidad matemática es difícil de modificar mediante el aprendizaje, su respuesta fue invertida mediante:

$$M_i^{Mat} = 5 - x_i, \quad (8)$$

donde x_i corresponde a la respuesta original del estudiante i y M_i^{Mat} a su valor recodificado. Posteriormente, esta medida fue estandarizada de acuerdo con la Ecuación 5, de manera que valores más altos representan una mayor creencia en que la capacidad matemática puede desarrollarse mediante el aprendizaje.

Por otra parte, se construyó un indicador de autoconcepto matemático a partir de las tres preguntas presentadas en el Anexo 11.3, relacionadas con la percepción del estudiante respecto de sus propias capacidades para aprender Matemática, su desempeño percibido y la facilidad que experimenta en esta asignatura. A diferencia de los indicadores de mentalidad, estas preguntas ya se encontraban formuladas de manera que valores mayores representaban una percepción más favorable de las propias capacidades, por lo que no fue necesario invertir sus respuestas.

El indicador se construyó mediante el promedio de las respuestas válidas:

$$A_i = \frac{1}{m_i} \sum_{j \in V_i} a_{ij}, \quad m_i \geq 2, \quad (9)$$

donde a_{ij} corresponde a la respuesta del estudiante i al ítem j , V_i representa el conjunto de respuestas válidas disponibles y m_i el número de respuestas utilizadas en el cálculo. Para construir el indicador se exigió contar con al menos dos de las tres respuestas disponibles, tras lo cual el promedio obtenido fue estandarizado mediante la Ecuación 5, de manera que valores más altos representan un mayor autoconcepto matemático.

5.2.3. Índice de condiciones laborales docentes

A diferencia de los indicadores anteriores, el índice de condiciones laborales docentes se construyó a partir de cuatro dimensiones medidas originalmente de distintas formas: estabilidad contractual, adecuación de la jornada laboral, proporción de horas no lectivas y cantidad de horas trabajadas fuera

de la jornada, cuyas preguntas de origen se presentan en el Anexo 11.4. A partir de esta información, cada componente fue definido de la siguiente manera:

- **Estabilidad contractual (E_i):** se asignó un valor de 2 a los contratos indefinidos, 1 a los contratos a plazo fijo y 0 a las restantes modalidades contractuales, de manera que valores mayores representaran una mayor estabilidad.
- **Adecuación de la jornada (A_i):** se asignó un valor de 0 a jornadas inferiores a 22 horas semanales, 2 a jornadas entre 22 y 45 horas y 1 a jornadas superiores a 45 y de hasta 60 horas semanales, considerando como situación más favorable el tramo intermedio.
- **Proporción de horas no lectivas (P_i):** se calculó como la razón entre las horas no lectivas y el total de horas contratadas:

$$P_i = \frac{H_i^{NL}}{H_i^T}, \quad (10)$$

donde H_i^{NL} representa las horas no lectivas y H_i^T el total de horas contratadas, por lo que valores mayores indican una mayor proporción de tiempo disponible para actividades distintas a la docencia directa.

- **Trabajo fuera de jornada (F_i):** corresponde al número de horas declaradas de trabajo realizado fuera de la jornada laboral. Como una mayor cantidad de estas horas representa una condición menos favorable, este componente fue incorporado posteriormente con signo inverso.

Dado que los cuatro componentes se encontraban medidos en escalas diferentes, cada uno fue estandarizado mediante la Ecuación 5 y posteriormente se calculó el índice como el promedio de los componentes válidos disponibles:

$$INDCOND_i = \frac{1}{m_i} \sum_{k \in V_i} C_{ik}, \quad m_i \geq 3, \quad (11)$$

donde C_{ik} corresponde a los componentes estandarizados $\{Z_{E_i}, Z_{A_i}, Z_{P_i}, -Z_{F_i}\}$, V_i representa el conjunto de componentes disponibles para el docente i y m_i el número de componentes válidos utilizados en su cálculo.

Finalmente, para construir el índice se exigió contar con información en al menos 3 de las 4 dimensiones, de manera que valores más altos de $INDCOND_i$ representan condiciones laborales relativamente más favorables dentro de la muestra analizada.

5.3 Unión de bases y construcción de la base final

Una vez limpias las distintas fuentes de información del SIMCE 2017, se procedió a su integración mediante un proceso de unión secuencial. La base de alumnos fue utilizada como archivo principal, al cual se incorporaron primero los cuestionarios de apoderados y estudiantes, ambos a nivel individual, y posteriormente los cuestionarios docentes de Matemática y Ciencias Naturales, vinculados a nivel de curso mediante las variables de identificación compartidas.

La Tabla 7 resume las principales pérdidas de información observadas al momento de realizar la consolidación de la base.

Tabla 7: Observaciones sin correspondencia al unir bases del SIMCE 2017

Fuente incorporada	Observaciones sin correspondencia
Cuestionario de apoderados	31.136
Cuestionario de estudiantes	2.974
Docente de Matemática	40.793
Docente de Ciencias Naturales	32.422

Como resultado de la integración de las distintas fuentes correspondientes al SIMCE 2017,

se obtuvo una base consolidada compuesta por 92.507 estudiantes. Esto representa una reducción de 107.325 observaciones respecto de la base de alumnos limpia de 2017, por lo que se conserva el 46,3 % de sus registros, mientras que, en comparación con las 247.495 observaciones disponibles inicialmente en la base de alumnos, la muestra corresponde al 37,4 %.

Sobre esta base se incorporaron posteriormente los puntajes obtenidos por los mismos estudiantes en el SIMCE 2013 de Matemática y Ciencias Naturales, junto con su promedio general de notas correspondiente al año 2016. Al requerir la disponibilidad de estos antecedentes académicos previos, la muestra se redujo de 92.507 a 73.796 estudiantes, lo que implica la pérdida adicional de 18.711 observaciones, equivalente al 20,2 % de la base consolidada de 2017. De esta manera, se obtuvo una base final de análisis compuesta por 73.796 estudiantes.

6 Estadísticas descriptivas

Una vez construida la base final de análisis, en esta sección se procede a caracterizar la muestra y examinar las principales variables utilizadas en el estudio, con el objetivo de obtener una primera aproximación a los patrones presentes en los datos antes de realizar las estimaciones econométricas.

6.1 Caracterización general de la muestra

Como punto de partida, la Tabla 8 presenta la composición general de la muestra final, mientras que las Figuras 1 y 2 muestran su distribución según el género de los estudiantes y de los docentes asociados a cada asignatura.

Tabla 8: Caracterización general de la muestra

Indicador	Valor
Número de estudiantes	73.796
Número de establecimientos	3.284
Número de cursos	4.525

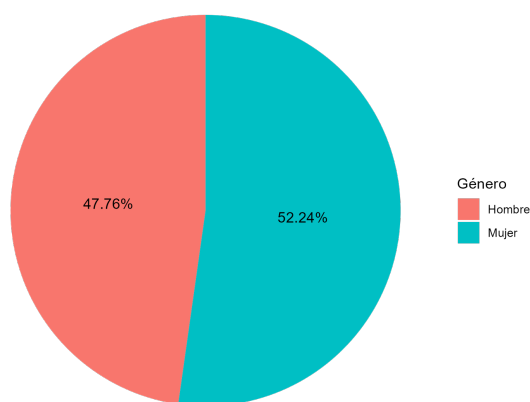


Figura 1: Distribución de estudiantes según género

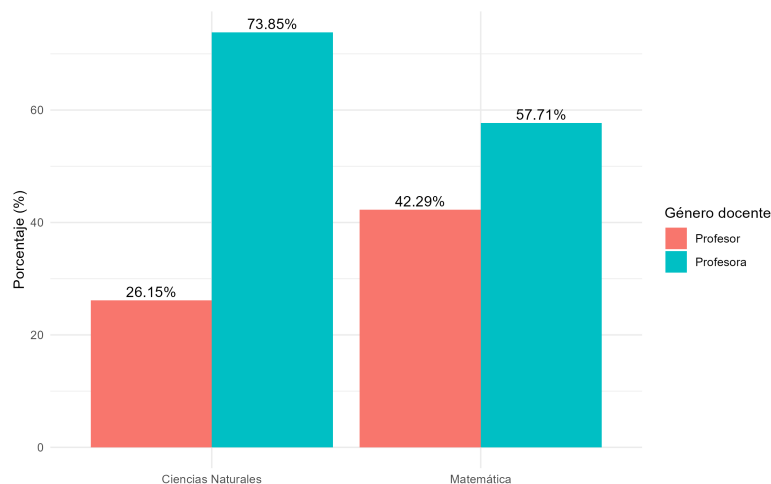


Figura 2: Distribución de estudiantes según género del docente y asignatura

La composición de estudiantes se encuentra relativamente equilibrada entre hombres y mujeres, mientras que, al considerar el género del docente asociado a cada estudiante, se observa una mayor presencia de profesoras en ambas asignaturas, particularmente en Ciencias Naturales. Cabe señalar que esta distribución no corresponde al número de docentes únicos, sino a la proporción de estudiantes que tiene un profesor o una profesora en cada asignatura.

6.2 Descripción de las principales variables

A partir de esta caracterización general, la Tabla 9 resume el comportamiento de las principales variables numéricas incorporadas posteriormente al análisis, incluyendo las medidas de rendimiento actual y previo, características de los docentes y los indicadores contruidos a partir de los cuestionarios.

Tabla 9: Estadísticos descriptivos de las principales variables numéricas

Variable	Media	DE	Rango
Puntaje Matemática 2017	268,54	49,06	134,27 – 399,42
Puntaje Ciencias Naturales 2017	265,44	46,99	142,46 – 417,36
Puntaje previo Matemática	265,20	49,52	82,87 – 381,26
Puntaje previo Ciencias Naturales	261,96	42,89	132,28 – 371,32
Promedio general 2016	5,12	1,75	1,0 – 6,9
Edad docente Matemática	40,44	11,67	22 – 78
Edad docente Ciencias Naturales	40,34	11,70	22 – 75
Experiencia docente Matemática	14,04	11,10	0 – 55
Experiencia docente Ciencias Naturales	14,12	11,04	0 – 52
Índice de Condiciones laborales Matemática	0,02	0,44	-3,13 – 1,86
Índice de Condiciones laborales Ciencias Naturales	0,00	0,46	-2,09 – 1,73
Índice de Mentalidad docente Matemática	0,03	0,97	-4,57 – 0,92
Índice de Mentalidad docente Ciencias Naturales	0,04	0,97	-4,57 – 0,89
Índice de Mentalidad general del estudiante	0,05	0,99	-2,25 – 1,95
Índice de Mentalidad hacia Matemática	0,07	0,98	-2,07 – 1,15

En términos generales, las características de los docentes presentan comportamientos similares entre Matemática y Ciencias Naturales, tanto en edad como en experiencia. Por su parte, los indicadores construidos se encuentran concentrados en torno a cero, en concordancia con los procedimientos de estandarización descritos previamente.

6.3 Diferencias descriptivas según género

Una vez caracterizada la muestra, resulta de particular interés observar cómo se distribuye el rendimiento según el género de los estudiantes, dado que estas diferencias constituyen el punto de partida para analizar posteriormente si su relación con el desempeño varía según el género del docente.

A partir de esta comparación, la Tabla 10 muestra una brecha favorable a los estudiantes hombres en ambas asignaturas, aunque esta resulta más marcada en Matemática que en Ciencias Naturales.

Tabla 10: Puntaje SIMCE promedio según género del estudiante

Asignatura	Género estudiante	Promedio	DE	N
Matemática	Hombre	270,80	48,94	35.244
Matemática	Mujer	266,48	49,07	38.552
Ciencias Naturales	Hombre	266,62	48,81	35.244
Ciencias Naturales	Mujer	264,36	45,24	38.552

Dado que el objetivo de la investigación no se limita a comparar el rendimiento de hombres y mujeres, sino que busca examinar su relación con el género del docente, el análisis incorpora ambas características de manera conjunta.

Para ello, las Tablas 11 y 12 presentan esta comparación para Matemática y Ciencias Naturales, respectivamente, permitiendo observar si las diferencias de rendimiento entre estudiantes varían según el género del docente que imparte la asignatura.

Tabla 11: Puntaje SIMCE en Matemática según género del docente y del estudiante

Género docente	Género estudiante	N	Promedio	DE
Profesor	Hombre	15.217	272,39	49,49
Profesor	Mujer	15.991	264,78	48,89
Profesora	Hombre	20.027	269,60	48,49
Profesora	Mujer	22.561	267,69	49,16

Acá se observa un patrón particularmente relevante para el análisis posterior, ya que la diferencia de rendimiento entre estudiantes hombres y mujeres es considerablemente menor cuando la asignatura es impartida por una profesora, lo que coincide con un mayor puntaje promedio de las estudiantes mujeres con profesoras y con el patrón contrario entre los estudiantes hombres.

Si bien estas comparaciones son únicamente descriptivas, permiten observar una posible variación en la relación entre el género del estudiante y el rendimiento según el género del docente, aspecto que posteriormente será evaluado mediante la interacción entre ambas variables.

Tabla 12: Puntaje SIMCE en Ciencias Naturales según género del docente y del estudiante

Género docente	Género estudiante	N	Promedio	DE
Profesor	Hombre	9.367	267,50	49,77
Profesor	Mujer	9.931	264,06	45,47
Profesora	Hombre	25.877	266,30	48,45
Profesora	Mujer	28.621	264,47	45,16

En Ciencias Naturales se mantiene una brecha favorable a los estudiantes hombres y esta también resulta menor cuando la docente es mujer, aunque las diferencias asociadas al género docente son más acotadas que las observadas en Matemática. De esta forma, la comparación entre ambas asignaturas sugiere que los patrones descriptivos vinculados al género del docente no presentan necesariamente la misma magnitud en todas las áreas evaluadas.

6.3.1. Prueba $t - test$

A partir de las diferencias descriptivas observadas entre hombres y mujeres, se aplicó una prueba t de diferencia de medias para evaluar si estas resultan estadísticamente significativas en Matemática y Ciencias Naturales.

En términos formales, la hipótesis nula plantea que el puntaje promedio de estudiantes hombres y mujeres es igual:

$$H_0 : \mu_H - \mu_M = 0, \quad (12)$$

mientras que la hipótesis alternativa establece que ambos promedios son distintos:

$$H_1 : \mu_H - \mu_M \neq 0. \quad (13)$$

Para realizar esta comparación se utilizó una prueba t de Welch para muestras independientes, la cual permite contrastar las medias de ambos grupos sin asumir igualdad de varianzas. Los resultados obtenidos para Matemática y Ciencias Naturales se presentan a continuación.

Código utilizado en R:

```
ttest_mate <- t.test(ptje_mate17 ~ gen_alu, data = base_final)

ttest_nat  <- t.test(ptje_nat17 ~ gen_alu, data = base_final)
```

Resultados:

```
> ttest_mate

Welch Two Sample t-test

t = 11.97, p-value < 2.2e-16

mean(Hombre) = 270.80
mean(Mujer)  = 266.48
IC 95% de la diferencia: [3.62, 5.03]

> ttest_nat

Welch Two Sample t-test

t = 6.48, p-value = 9.257e-11

mean(Hombre) = 266.62
mean(Mujer)  = 264.36
IC 95% de la diferencia: [1.57, 2.93]
```

Los resultados indican que las diferencias de puntaje entre hombres y mujeres son estadística-

mente significativas ¹, lo que implica que las brechas observadas difícilmente pueden atribuirse únicamente a variaciones aleatorias de la muestra. En consecuencia, se rechaza la hipótesis nula de igualdad de medias en ambas asignaturas, manteniéndose una brecha favorable a los estudiantes hombres y de mayor magnitud en Matemática que en Ciencias Naturales.

Sin embargo, dado que estas comparaciones consideran únicamente las diferencias de medias entre ambos grupos, sin incorporar otras características contextuales, deben interpretarse como evidencia descriptiva previa a las estimaciones econométricas.

6.4 Características familiares

Finalmente, la caracterización de la muestra se complementa con algunos antecedentes del entorno familiar que posteriormente son incorporados como variables de control.

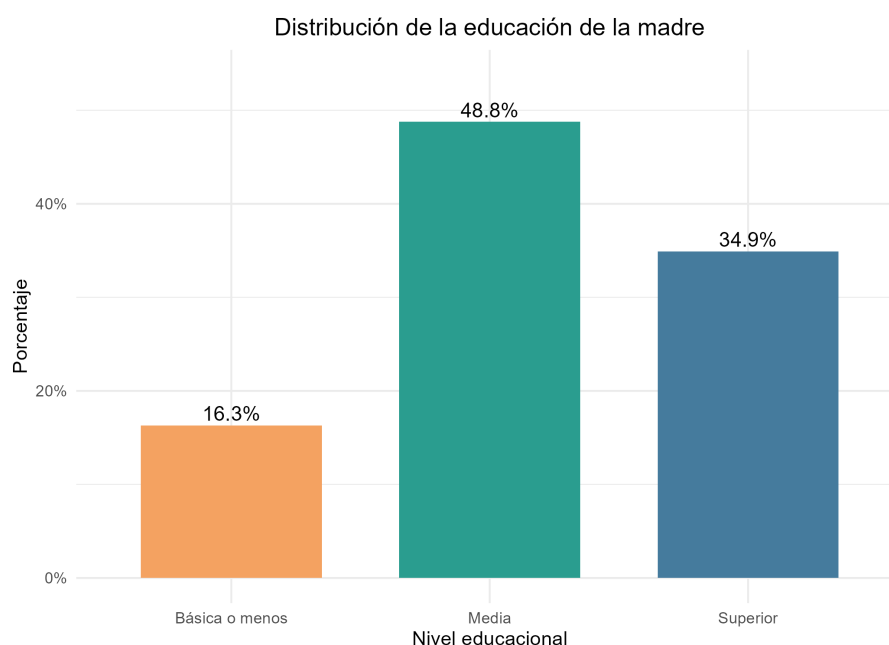


Figura 3: Distribución de la educación de la madre

¹El estadístico t muestra cuán grande es la diferencia observada entre ambos grupos en relación con la variabilidad de los datos, por ende, mientras más se aleja de cero, mayor es la evidencia de que los promedios son distintos, por su parte, el p -value indica qué tan compatible es el resultado observado con la hipótesis nula, por lo que valores menores a 0,05 sugieren que esa igualdad es poco probable y permiten rechazarla. Finalmente, el intervalo de confianza al 95 % indica el rango dentro del cual probablemente se encuentra la diferencia real entre los promedios; si todo ese rango se encuentra por encima o por debajo de cero, se considera que existe una diferencia estadísticamente significativa.

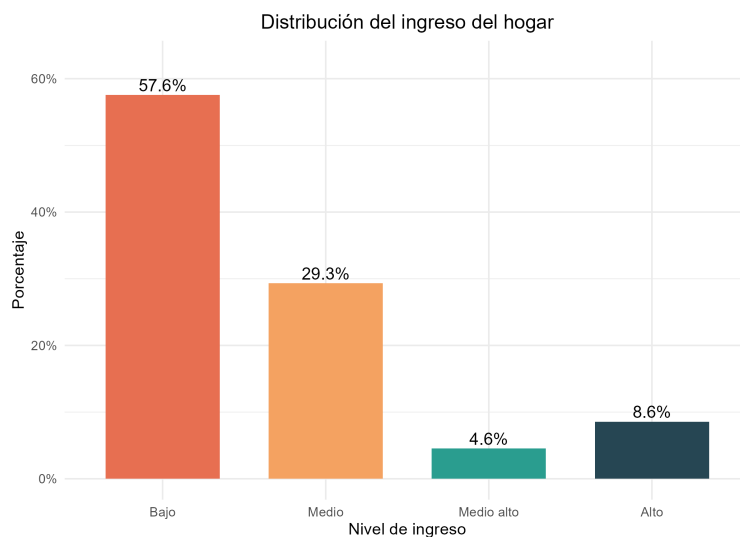


Figura 4: Distribución del ingreso del hogar

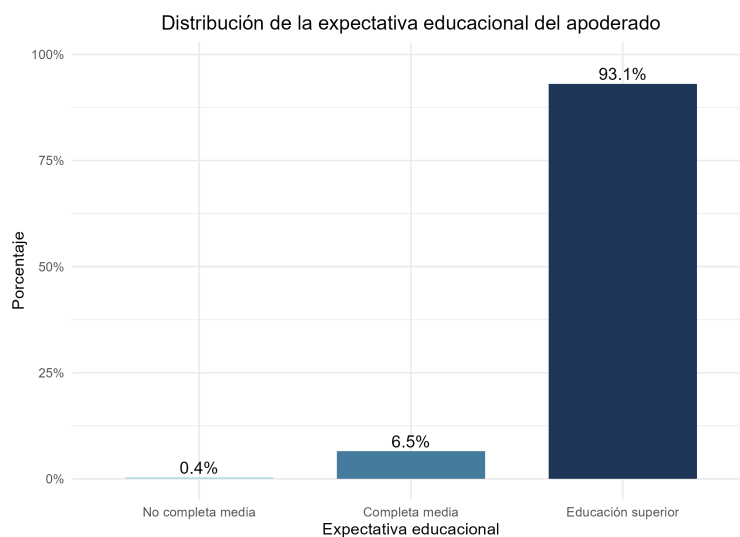


Figura 5: Distribución de la expectativa educacional del apoderado

En conjunto, las estadísticas descriptivas muestran una muestra en la que existen diferencias de rendimiento entre hombres y mujeres, junto con una mayor presencia de profesoras en ambas asignaturas. A ello se suma un entorno caracterizado principalmente por madres con educación media, hogares concentrados en niveles de ingreso bajos y medios y altas expectativas educacionales, antecedentes que serán considerados en la estrategia metodológica desarrollada a continuación.

7 Metodología

A partir de los antecedentes desarrollados en las secciones anteriores y de la base de datos construida para el análisis, se presenta la estrategia metodológica utilizada en esta investigación, la cual se desarrolla en dos etapas.²

En primer lugar, se estima la relación entre el género del docente y el rendimiento de los estudiantes mediante una serie de especificaciones que incorporan progresivamente características académicas, familiares y docentes, para posteriormente introducir efectos fijos que permiten controlar por fuentes adicionales de heterogeneidad no observada.

En segundo lugar, se desarrollan estrategias de análisis complementarias orientadas a evaluar si los patrones encontrados son consistentes con los mecanismos de modelo de rol y sesgo docente planteados en el marco teórico.

7.1 Estimación del efecto del género docente sobre el rendimiento estudiantil

Esta primera etapa considera seis especificaciones que buscan evaluar la relación entre el género del docente y el rendimiento académico bajo distintos conjuntos de controles y estructuras de efectos fijos.

Las primeras cuatro se estiman separadamente para Matemática y Ciencias Naturales e incorporan progresivamente antecedentes académicos, familiares y docentes, con el objetivo de observar cómo evoluciona la relación de interés a medida que se consideran factores que también podrían estar asociados al rendimiento y, al mismo tiempo, evaluar la estabilidad del coeficiente principal frente a distintas especificaciones.

²En todas las especificaciones, los errores estándar se agrupan a nivel de establecimiento (RBD), permitiendo considerar la posible correlación entre las observaciones pertenecientes a un mismo colegio.

7.1.1. Modelo 1: control por rendimiento académico previo

Como primera aproximación, se estima un modelo lineal que incorpora el puntaje obtenido por el estudiante en el SIMCE 2013 y el promedio general de notas correspondiente al año 2016. La incorporación de estas variables tiene como objetivo comparar estudiantes considerando diferencias previas en su trayectoria académica y, con ello, reducir la posibilidad de que la relación observada con el género docente refleje únicamente diferencias de rendimiento existentes con anterioridad.

Para cada asignatura s , y tomando como categorías de referencia a los estudiantes hombres y a los docentes hombres, la especificación corresponde a:

$$Y_{is}^z = \beta_0 + \beta_1 \text{Mujer}_i + \beta_2 \text{Profesora}_{is} + \beta_3 (\text{Mujer}_i \times \text{Profesora}_{is}) + \rho \text{PuntajePrevio}_{is}^z + \delta \text{Promedio}_i^z + \varepsilon_{is}. \quad (14)$$

donde:

- Y_{is}^z corresponde al puntaje SIMCE 2017 estandarizado del estudiante i en la asignatura s .
- Mujer_i es una variable que toma el valor 1 cuando el estudiante es mujer y 0 cuando es hombre.
- Profesora_{is} es una variable que toma el valor 1 cuando el estudiante i tiene una profesora en la asignatura s y 0 cuando tiene un profesor.
- $\text{Mujer}_i \times \text{Profesora}_{is}$ corresponde a la interacción entre el género del estudiante y el género del docente.
- $\text{PuntajePrevio}_{is}^z$ corresponde al puntaje SIMCE 2013 estandarizado obtenido por el estudiante i en la asignatura s .
- Promedio_i^z corresponde al promedio general de notas del estudiante i en 2016, estandarizado.

- ε_{is} corresponde al término de error para cada estudiante i y asignatura s .

Dadas las categorías de referencia definidas, β_1 representa la diferencia de rendimiento entre estudiantes mujeres y hombres cuando el docente es hombre, mientras que β_2 recoge la diferencia asociada a tener una profesora para los estudiantes hombres. Por su parte, β_3 , correspondiente a la interacción entre ambas variables, indica cuánto cambia la asociación entre tener una profesora y el rendimiento para las estudiantes mujeres en comparación con los hombres, constituyendo así el principal coeficiente de interés.

En consecuencia, la asociación total entre tener una profesora y el rendimiento de las estudiantes mujeres viene dada por:

$$\beta_2 + \beta_3 \tag{15}$$

Cabe destacar que, dado que la variable dependiente se encuentra estandarizada, los coeficientes pueden interpretarse en términos de desviaciones estándar del puntaje SIMCE.

Por otro lado, si bien esta primera especificación permite considerar diferencias previas en el desempeño académico de los estudiantes, estas no representan la totalidad de los factores que pueden encontrarse asociados al rendimiento, ya que estudiantes con trayectorias académicas similares pueden pertenecer a contextos familiares diferentes, cuyas características podrían influir tanto en sus resultados educativos como en el entorno escolar en el que se desenvuelven. Por esta razón, la siguiente especificación amplía el conjunto de controles incorporando antecedentes del hogar.

7.1.2. Modelo 2: incorporación de características familiares

Manteniendo las variables consideradas en el modelo anterior, en este segundo modelo se incorpora el nivel educacional de la madre, la disponibilidad de información sobre la educación del padre, el ingreso mensual del hogar y las expectativas educacionales declaradas por el apoderado.

La educación de la madre ($EducMadre_i$) se incorpora mediante tres categorías: educación básica o menos, educación media y educación superior, utilizando la primera como categoría de referencia.

Por su parte, la educación del padre ($Padre_i$) se incluye mediante el indicador de disponibilidad de información construido previamente, mientras que las expectativas educacionales ($Expectativa_i$) del apoderado distinguen entre no completar la educación media, completar la educación media y alcanzar educación superior, tomando la primera como referencia.

Finalmente, el ingreso mensual ($Ingreso_i$) del hogar se incorpora mediante cuatro categorías: ingreso bajo, medio, medio alto y alto, utilizando el nivel de ingreso bajo como categoría de referencia.

3

De esta manera, para cada asignatura s , el modelo se expresa como:

$$Y_{is}^z = \beta_0 + \beta_1 Mujer_i + \beta_2 Profesora_{is} + \beta_3 (Mujer_i \times Profesora_{is}) + \rho PuntajePrevio_{is}^z + \delta Promedio_i^z + \gamma' F_i + \varepsilon_{is}, \quad (16)$$

donde F_i representa el conjunto de características familiares incorporadas en esta especificación y γ' el conjunto de coeficientes asociados a estas características.

La inclusión de estos controles permite evaluar si la relación entre el género del docente y el

³Como prueba de robustez, se estimó alternativamente la especificación utilizando los 15 tramos de ingreso originales disponibles en el cuestionario de apoderados. Los resultados se presentan en el Anexo 11.5 y muestran que las conclusiones del modelo se mantienen al utilizar una categorización más desagregada.

rendimiento, particularmente aquella capturada por la interacción $Mujer_i \times Profesora_{is}$, se mantiene una vez consideradas diferencias observables en el entorno familiar de los estudiantes. Sin embargo, aun después de incorporar estos antecedentes, la estimación podría seguir recogiendo diferencias asociadas a características propias del docente que también se relacionen con el desempeño estudiantil. Por esta razón, la siguiente especificación amplía nuevamente el conjunto de controles incorporando estas características.

7.1.3. Modelo 3: incorporación de características docentes

Manteniendo los controles de las especificaciones anteriores, el tercer modelo añade características observables del docente que pueden encontrarse asociadas a la forma en que se desarrolla la enseñanza, a la motivación de los estudiantes y a la capacidad de responder a sus necesidades de aprendizaje, por lo que parte de la relación atribuida inicialmente al género docente podría estar recogiendo también estas diferencias.

En particular, se incorporan las siguientes características docentes:

- $Titulo_{is}$: indica si el docente cuenta con título profesional.
- $Especializacion_{is}$: identifica si el docente posee especialización en la asignatura correspondiente.
- $Experiencia_{is}^z$: corresponde a los años de experiencia del profesor, estandarizados. Esta variable se incorpora tanto de forma lineal como cuadrática, permitiendo que su relación con el rendimiento no sea necesariamente constante a medida que aumentan los años de ejercicio profesional.
- $Condiciones_{is}$: corresponde al índice de condiciones laborales docentes construido previamente.

- *MentalidadProf_{is}*: corresponde al indicador de mentalidad del profesor construido previamente.

De esta manera, para cada asignatura s , el modelo se expresa como:

$$\begin{aligned} Y_{is}^z = & \beta_0 + \beta_1 \text{Mujer}_i + \beta_2 \text{Profesora}_{is} + \beta_3 (\text{Mujer}_i \times \text{Profesora}_{is}) \\ & + \rho \text{PuntajePrevio}_{is}^z + \delta \text{Promedio}_i^z + \gamma' F_i \\ & + \theta' D_{is} + \varepsilon_{is}, \end{aligned} \tag{17}$$

donde D_{is} representa el conjunto de características docentes añadidas en esta especificación y θ' el conjunto de coeficientes asociados a ellas.

La incorporación de estos controles permite evaluar si la relación entre el género del docente y el rendimiento se mantiene una vez consideradas diferencias observables entre profesores y profesoras.

Hasta este punto, las especificaciones anteriores han incorporado progresivamente características observables asociadas a los estudiantes, sus familias y sus docentes. Sin embargo, aun considerando estos antecedentes, pueden persistir diferencias no observadas que se encuentren relacionadas tanto con el rendimiento de los estudiantes como con la asignación de docentes, lo que podría afectar la estimación de la relación de interés.

Frente a esta preocupación, a partir de la siguiente especificación se comienza a incorporar una estrategia de efectos fijos, tomando como referencia la estrategia empírica desarrollada por Paredes (2014).

En base a esto, como primer paso se introducen efectos fijos por establecimiento, con el propósito de controlar aquellas características, observadas y no observadas, que son comunes a los estudiantes de un mismo colegio y que, por tanto, no varían entre ellos (como disponibilidad de recursos, infraestructura, prácticas de gestión, entre otras).

7.1.4. Modelo 4: efectos fijos por establecimiento

A diferencia de las especificaciones anteriores, este modelo incorpora un efecto fijo para cada establecimiento, manteniendo los controles académicos, familiares y docentes previamente considerados. De esta manera, las características que son comunes a todos los estudiantes de un mismo colegio quedan absorbidas por el efecto fijo correspondiente y dejan de contribuir a la estimación de los coeficientes de interés.

La especificación se expresa como:

$$Y_{is}^z = \alpha_r + \beta_1 \text{Mujer}_i + \beta_2 \text{Profesora}_{is} + \beta_3 (\text{Mujer}_i \times \text{Profesora}_{is}) + \rho \text{PuntajePrevio}_{is}^z + \delta \text{Promedio}_i^z + \gamma' F_i + \theta' D_{is} + \varepsilon_{isr}, \quad (18)$$

donde α_r corresponde al efecto fijo del establecimiento r , F_i mantiene el conjunto de características familiares definido en el Modelo 2 y D_{is} representa las características docentes incorporadas en el Modelo 3.

La principal diferencia respecto de las especificaciones anteriores radica en que la relación de interés pasa a estimarse a partir de las diferencias observadas entre estudiantes pertenecientes a un mismo establecimiento. En consecuencia, las características propias de los colegios, tanto aquellas que pueden observarse directamente como aquellas para las que no se dispone de información, dejan de influir en la estimación, permitiendo evaluar si el género del docente continúa asociado al rendimiento de los estudiantes aun cuando la comparación se realiza entre alumnos pertenecientes a un mismo colegio.

Sin embargo, aun dentro de un mismo establecimiento pueden existir características no observadas de los estudiantes, como ciertas habilidades, motivaciones u otros atributos persistentes, que se relacionen con su rendimiento y que no sean consideradas mediante los controles anteriores. Dado que

los efectos fijos por establecimiento no permiten controlar estas diferencias individuales, la siguiente especificación traslada la estrategia de efectos fijos al nivel del estudiante.

7.1.5. Modelo 5: efectos fijos por estudiante

Para estimar esta especificación, la base de datos se reorganizó de tal manera que cada estudiante aporta dos observaciones, una correspondiente a Matemática y otra a Ciencias Naturales, lo que permite aprovechar la variación existente dentro del mismo alumno entre ambas asignaturas, incluyendo posibles cambios en el género del docente.

Al incorporar efectos fijos por estudiante, aquellas características que permanecen constantes entre ambas observaciones, como el género del alumno, las características familiares y el promedio general de notas, quedan absorbidas por el efecto fijo individual y dejan de estimarse de manera separada. En cambio, las características que presentan variación entre Matemática y Ciencias Naturales pueden continuar contribuyendo a la estimación, entre ellas la interacción entre el género del estudiante y el género del docente.

En particular, el modelo se expresa como:

$$\begin{aligned} Y_{is}^z = & \alpha_i + \lambda_s + \beta_2 Profesora_{is} + \beta_3 (Mujer_i \times Profesora_{is}) \\ & + \rho PuntajePrevio_{is}^z + \kappa (Mujer_i \times PuntajePrevio_{is}^z) \\ & + \theta' D_{is} + \phi (Mujer_i \times Naturales_s) + \varepsilon_{is}, \end{aligned} \quad (19)$$

donde, además de las variables definidas en las especificaciones anteriores:

- α_i corresponde al efecto fijo del estudiante, el cual absorbe las características observadas y no observadas que permanecen constantes para un mismo alumno entre Matemática y Ciencias Naturales.

- λ_s corresponde al efecto fijo de asignatura y permite considerar diferencias generales entre Matemática y Ciencias Naturales que afectan de manera común a los estudiantes, como aquellas asociadas a los contenidos evaluados, las características de cada prueba o las particularidades propias de cada área.
- $Mujer_i \times PuntajePrevio_{is}^z$ permite que la asociación entre el rendimiento previo y el rendimiento actual pueda diferir entre hombres y mujeres, siendo κ el coeficiente asociado a esta diferencia
- $Mujer_i \times Naturales_s$ se incorpora porque, a diferencia de los modelos anteriores, Matemática y Ciencias Naturales se estiman conjuntamente y el género del estudiante queda absorbido por α_i , por lo que esta interacción permite observar cómo varía la diferencia de rendimiento entre ambas asignaturas para hombres y mujeres.

Por su parte, D_{is} contiene las características docentes que pueden variar entre ambas asignaturas, correspondientes a la titulación, especialización y experiencia del profesor, incorporándose esta última tanto de forma lineal como cuadrática, mientras que θ' representa el conjunto de coeficientes asociados a ellas.

Bajo esta especificación, β_3 continúa siendo el principal coeficiente de interés y mantiene la interpretación presentada anteriormente.

Sin embargo, aun después de incorporar efectos fijos por estudiante y asignatura, pueden permanecer características propias de la combinación entre curso y asignatura que son compartidas por todos los estudiantes que reciben esa clase. Estas pueden incluir, por ejemplo, características no observadas del docente, las cuales no son capturadas por los efectos fijos anteriores. Por esta razón, la siguiente especificación incorpora efectos fijos por curso-asignatura, permitiendo controlar también

aquellos factores comunes a los estudiantes que reciben una determinada asignatura dentro del mismo curso.

7.1.6. Modelo 6: efectos fijos por curso-asignatura

Como última especificación de esta primera etapa, se incorporan efectos fijos para cada combinación entre establecimiento, curso y asignatura mediante la construcción de un identificador único a partir del RBD, el identificador del curso y la asignatura correspondiente, de modo que cada grupo reúne a los estudiantes que comparten una misma clase de Matemática o Ciencias Naturales dentro de un establecimiento.

En particular, el modelo se expresa como:

$$\begin{aligned} Y_{is}^z = & \alpha_i + \eta_{rcs} + \beta_3 (Mujer_i \times Profesora_{is}) \\ & + \rho PuntajePrevio_{is}^z + \kappa (Mujer_i \times PuntajePrevio_{is}^z) \\ & + \phi (Mujer_i \times Naturales_s) + \varepsilon_{is}, \end{aligned} \quad (20)$$

donde, además de las variables definidas anteriormente:

- η_{rcs} corresponde al efecto fijo asociado a cada combinación entre establecimiento r , curso c y asignatura s , permitiendo absorber aquellas características comunes a todos los estudiantes que reciben una determinada asignatura dentro del mismo curso.

A diferencia del Modelo 5, la variable $Profesora_{is}$ y las demás características docentes dejan de estimarse de manera separada, ya que presentan el mismo valor para los estudiantes que comparten una determinada combinación curso-asignatura y, por tanto, quedan absorbidas por η_{rcs} . Sin embargo, la interacción $Mujer_i \times Profesora_{is}$ continúa siendo estimable, debido a que el género del estudiante sí varía entre los alumnos que comparten una misma clase.

De esta forma, β_3 continúa siendo el principal coeficiente de interés y permite evaluar si, dentro de una misma combinación curso-asignatura, la asociación entre tener una profesora y el rendimiento difiere entre estudiantes mujeres y hombres.

Esta especificación constituye el modelo más restrictivo de la primera etapa del análisis, al combinar efectos fijos por estudiante con efectos fijos por curso y asignatura, lo que permite evaluar si el patrón asociado al género del docente se mantiene una vez controladas tanto las características persistentes de los estudiantes como aquellas compartidas por quienes reciben una misma asignatura dentro del mismo curso.

7.2 Estrategias para evaluar mecanismos subyacentes

Una vez definida la estrategia utilizada para estimar la relación entre el género del docente y el rendimiento estudiantil, esta sección presenta la segunda etapa de la metodología, orientada a examinar los posibles mecanismos subyacentes a esta relación: el modelo de rol y el sesgo docente.

Para ello, se plantean distintas estrategias complementarias que permiten evaluar si los patrones observados son consistentes con las predicciones asociadas a cada mecanismo, sin que ello implique una identificación directa de su existencia.

7.2.1. Estrategias asociadas al sesgo docente

El primer grupo de estrategias busca examinar si los patrones asociados al género del docente pueden ser consistentes con diferencias en el comportamiento de profesores y profesoras hacia sus estudiantes según su género. Para ello, se consideran las aproximaciones mostradas a continuación.

7.2.1.1. Expectativas académicas del docente

Esta estrategia se inspira en el análisis desarrollado por Paredes (2014) y busca evaluar si las expectati-

vas académicas de profesores y profesoras responden de manera diferente a la composición de género del curso. En este sentido, si existen diferencias en el comportamiento docente asociadas al género de los estudiantes, cabría esperar que la relación entre la proporción de hombres presentes en el curso y las expectativas académicas varíe según el género del docente.

Para implementar este análisis, se construye una base cuya unidad de observación corresponde a cada combinación entre curso y asignatura, considerando únicamente cursos mixtos para que exista variación en la proporción de estudiantes hombres.

La especificación se expresa como:

$$\begin{aligned} Expectativa_{rcs} = & \alpha_0 + \alpha_1 Profesora_{rcs} + \alpha_2 PropHombres_{rc} \\ & + \alpha_3 (Profesora_{rcs} \times PropHombres_{rc}) + \alpha_4 RendPrevio_{rcs} + u_{rcs}, \end{aligned} \quad (21)$$

donde:

- $Expectativa_{rcs}$ corresponde a la expectativa educacional declarada por el docente en seis categorías para el curso c , establecimiento r y asignatura s .
- $PropHombres_{rc}$ representa la proporción de estudiantes hombres presentes en el curso.
- $RendPrevio_{rcs}$ corresponde al rendimiento previo promedio de los estudiantes del curso en la asignatura correspondiente.
- $Profesora_{rcs} \times PropHombres_{rc}$ corresponde a la interacción de interés, cuyo coeficiente α_3 permite evaluar si la relación entre la composición de género del curso y las expectativas académicas difiere entre profesores y profesoras.

De esta manera, un valor de α_3 estadísticamente distinto de cero indicaría que las expectativas de profesores y profesoras responden de manera diferente a medida que cambia la proporción de es-

tudiantes hombres del curso. Por el contrario, un coeficiente igual a cero indicaría que no se observa evidencia de una respuesta diferencial según el género del docente.

Finalmente, dado que las expectativas docentes se encuentran medidas mediante seis categorías ordenadas, se estima adicionalmente un modelo *ordered probit* como prueba de robustez. Este modelo permite respetar el orden existente entre las categorías sin asumir que la distancia entre cada una de ellas sea necesariamente la misma, permitiendo comprobar si las conclusiones se mantienen bajo una forma alternativa de tratar esta variable.

7.2.1.2. Composición de género del curso

Como segunda aproximación, se evalúa si la asociación entre tener una profesora y el rendimiento de las estudiantes varía según la composición de género del curso. Esta estrategia toma como referencia la predicción planteada por Paredes (2014), según la cual, si el patrón observado estuviera explicado principalmente por diferencias en el trato de profesores y profesoras hacia estudiantes hombres y mujeres dentro del aula, cabría esperar que la asociación con una profesora fuese mayor en cursos mixtos y disminuyera cuando las estudiantes se encuentran en cursos compuestos únicamente por mujeres.

Para implementar esta estrategia, los cursos se clasifican según su composición de género y el análisis se concentra en las estudiantes pertenecientes a cursos mixtos y a cursos compuestos únicamente por mujeres, ya que la comparación busca evaluar si la asociación entre tener una profesora y el rendimiento de las alumnas cambia cuando estas comparten el aula con estudiantes hombres.

La especificación se expresa como:

$$\begin{aligned} Y_{is}^z = & \alpha_i + \lambda_s + \beta_1 Profesora_{is} + \beta_2 (Profesora_{is} \times SoloMujeres_i) \\ & + \rho PuntajePrevio_{is}^z + \theta' D_{is} + \varepsilon_{is}, \end{aligned} \quad (22)$$

donde, además de las variables definidas previamente:

- $SoloMujeres_i$ identifica a las estudiantes pertenecientes a cursos compuestos exclusivamente por mujeres, siendo los cursos mixtos la categoría de referencia.
- $Profesora_{is} \times SoloMujeres_i$ corresponde a la interacción de interés, cuyo coeficiente β_2 indica cuánto cambia la asociación entre tener una profesora y el rendimiento cuando la estudiante pertenece a un curso compuesto exclusivamente por mujeres respecto de un curso mixto.

De esta manera, β_1 representa la asociación entre tener una profesora y el rendimiento para las estudiantes pertenecientes a cursos mixtos, mientras que para aquellas que se encuentran en cursos solo de mujeres esta asociación corresponde a:

$$\beta_1 + \beta_2. \quad (23)$$

Bajo la predicción asociada al sesgo docente evaluada mediante esta estrategia, se esperaría que la asociación con una profesora fuese menor en los cursos compuestos exclusivamente por mujeres que en los cursos mixtos. Aun así, un patrón diferente no permitiría descartar otras posibles formas de sesgo docente, sino únicamente señalar que los resultados no serían consistentes con esta predicción específica, según la cual el efecto sería más pronunciado cuando hombres y mujeres comparten el aula.

7.2.2. Estrategias asociadas al modelo de rol

El segundo grupo de estrategias busca examinar si los patrones asociados al género del docente pueden ser consistentes con un mecanismo de modelo de rol, según el cual la presencia de una profesora podría actuar como referente para las estudiantes e influir en su rendimiento, percepciones o

creencias respecto de sus propias capacidades. Para ello, se consideran las aproximaciones presentadas a continuación.

7.2.2.1. Diferencias entre asignaturas

La primera estrategia toma como referencia la lógica desarrollada por Paredes (2014) y busca evaluar si la asociación entre tener una profesora y el rendimiento de las estudiantes presenta diferencias entre Matemática y Ciencias Naturales. La idea es que, si la presencia de una docente mujer adquiere mayor relevancia como referente en áreas donde existen mayores estereotipos de género, cabría esperar que dicha asociación no fuese necesariamente igual en ambas asignaturas.

Para implementar este análisis, la muestra se restringe a las estudiantes mujeres y se utilizan conjuntamente sus observaciones de Matemática y Ciencias Naturales, manteniendo la estructura de efectos fijos por estudiante utilizada previamente. A partir de esta especificación, se incorpora una interacción entre tener una profesora y la asignatura de Ciencias Naturales, lo que permite evaluar si la asociación entre el género docente y el rendimiento de las estudiantes difiere entre ambas áreas.

La especificación se expresa como:

$$Y_{is}^z = \alpha_i + \lambda_s + \beta_1 \text{Profesora}_{is} + \beta_2 (\text{Profesora}_{is} \times \text{Naturales}_s) + \rho \text{PuntajePrevio}_{is}^z + \theta' D_{is} + \varepsilon_{is}, \quad (24)$$

donde, además de las variables definidas anteriormente:

- Naturales_s identifica las observaciones correspondientes a Ciencias Naturales, siendo Matemática la categoría de referencia.
- $\text{Profesora}_{is} \times \text{Naturales}_s$ corresponde a la interacción de interés, cuyo coeficiente β_2 indica cuánto cambia la asociación entre tener una profesora y el rendimiento en Ciencias Naturales

respecto de Matemática.

En esta especificación, β_1 representa la asociación entre tener una profesora y el rendimiento de las estudiantes en Matemática, mientras que en Ciencias Naturales dicha asociación corresponde a:

$$\beta_1 + \beta_2. \quad (25)$$

De esta manera, la comparación entre ambas asignaturas permite evaluar si la asociación con una profesora adquiere distinta magnitud según el área en que se desempeña la estudiante, lo que aporta evidencia complementaria para examinar si los resultados son consistentes con el mecanismo de modelo de rol.

7.2.2.2. Educación de la madre

La segunda estrategia busca evaluar si la asociación entre tener una profesora y el rendimiento de las estudiantes varía según el nivel educacional de la madre. En este caso, la educación materna se utiliza como una aproximación a la presencia de referentes femeninos académicos en el entorno cercano de la estudiante, bajo la idea de que una docente mujer podría adquirir mayor relevancia como modelo de rol cuando estos referentes son menos frecuentes. Por esta razón, el análisis se concentra en las estudiantes mujeres.

Para examinar este patrón, primero se estima la especificación por separado para cada nivel de educación materna, permitiendo comparar la asociación entre tener una profesora y el rendimiento entre los distintos grupos. Posteriormente, se utiliza una estimación conjunta con interacciones para evaluar formalmente si las diferencias observadas entre ellos son estadísticamente significativas.

Considerando la educación básica o menos como categoría de referencia, la especificación conjunta se expresa como:

$$\begin{aligned} Y_{is}^z = & \alpha_i + \lambda_s + \beta_1 Profesora_{is} \\ & + \beta_2 (Profesora_{is} \times MadreMedia_i) + \beta_3 (Profesora_{is} \times MadreSuperior_i) \quad (26) \\ & + \rho PuntajePrevio_{is}^z + \theta' D_{is} + \varepsilon_{is}, \end{aligned}$$

donde, además de las variables definidas anteriormente:

- *MadreMedia_i* y *MadreSuperior_i* identifican a las estudiantes cuyas madres presentan educación media y educación superior, respectivamente.
- *Profesora_{is} × MadreMedia_i* y *Profesora_{is} × MadreSuperior_i* permiten evaluar si la asociación entre tener una profesora y el rendimiento cambia según el nivel educacional de la madre, siendo β_2 y β_3 las diferencias respecto del grupo de referencia.

Debido a que el nivel educacional de la madre permanece constante para una misma estudiante entre Matemática y Ciencias Naturales, su efecto directo queda absorbido por el efecto fijo individual; sin embargo, sus interacciones con *Profesora_{is}* pueden estimarse porque el género del docente puede variar entre ambas asignaturas.

7.2.2.3. Autoconcepto matemático

La siguiente estrategia busca evaluar si tener una profesora de Matemática se encuentra asociado con una percepción más favorable de las propias capacidades matemáticas de las estudiantes. Esta aproximación se relaciona con el mecanismo de modelo de rol, en la medida en que la presencia de una docente mujer podría contribuir a reforzar la percepción que las alumnas tienen sobre su capacidad para desenvolverse en esta área.

Para este análisis se utiliza como variable dependiente el índice estandarizado de autoconcepto matemático construido previamente y, debido a que este indicador se observa una única vez por es-

tudiante, no es posible mantener la estructura de efectos fijos individuales utilizada en las estrategias anteriores, por lo que la estimación incorpora efectos fijos por establecimiento.

La especificación se expresa como:

$$\begin{aligned} \text{Autoconcepto}_i^z = & \alpha_r + \beta_1 \text{Mujer}_i + \beta_2 \text{Profesora}_i + \beta_3 (\text{Mujer}_i \times \text{Profesora}_i) \\ & + \rho \text{PuntajePrevio}_i^z + \delta \text{Promedio}_i^z + \theta' D_i + \varepsilon_i, \end{aligned} \quad (27)$$

donde, además de las variables definidas anteriormente:

- Autoconcepto_i^z corresponde al índice estandarizado de autoconcepto matemático del estudiante i , donde valores mayores representan una percepción más favorable de sus propias capacidades en Matemática.
- α_r corresponde al efecto fijo del establecimiento, permitiendo controlar aquellas características comunes a los estudiantes pertenecientes a un mismo colegio.

La interacción $\text{Mujer}_i \times \text{Profesora}_i$ mantiene la lógica de interpretación utilizada en las especificaciones anteriores, pero en este caso permite observar si la asociación entre tener una profesora de Matemática y el autoconcepto matemático difiere entre estudiantes mujeres y hombres.

7.2.2.4. Mentalidad hacia Matemática

De manera complementaria al análisis de autoconcepto matemático, esta estrategia evalúa si el género del docente se encuentra asociado con las creencias de los estudiantes respecto de la posibilidad de desarrollar sus capacidades matemáticas mediante el aprendizaje.

La especificación se expresa como:

$$\begin{aligned} \text{MentalidadMate}_i^z = & \alpha_r + \beta_1 \text{Mujer}_i + \beta_2 \text{Profesora}_i + \beta_3 (\text{Mujer}_i \times \text{Profesora}_i) \\ & + \rho \text{PuntajePrevio}_i^z + \delta \text{Promedio}_i^z + \theta' D_i + \varepsilon_i, \end{aligned} \quad (28)$$

donde $MentalidadMate_i^z$ corresponde al indicador estandarizado de mentalidad hacia Matemática construido previamente y la interacción $Mujer_i \times Profesora_i$ permite evaluar si la asociación entre tener una profesora de Matemática y estas creencias difiere entre estudiantes mujeres y hombres, manteniéndose sin cambios el resto de la estructura utilizada en el análisis de autoconcepto matemático.

De manera complementaria, esta misma especificación se estima utilizando el indicador de mentalidad general, con el propósito de observar si los patrones encontrados se concentran en el ámbito matemático o también se presentan en creencias más generales respecto del aprendizaje.

8 Resultados

8.1 Relación entre el género docente y el rendimiento estudiantil

8.1.1. Resultados con controles observables

En primer lugar, se analizan las tres especificaciones que incorporan progresivamente características observables, correspondientes a los Modelos 1, 2 y 3. Para esto, la Tabla 13 presenta los coeficientes directamente relacionados con el género del estudiante y del docente, junto con la asociación estimada de tener una profesora para las estudiantes mujeres y las principales medidas de ajuste de cada especificación, mientras que los resultados completos de las estimaciones se presentan en el Anexo 11.6.

Tabla 13: Resultados de los modelos con controles observables

	Matemática			Ciencias Naturales		
Coefficientes	Modelo 1	Modelo 2	Modelo 3	Modelo 1	Modelo 2	Modelo 3
Mujer	-0,0889***	-0,0844***	-0,0842***	0,0221	0,0171	0,0148
Error estándar	(0,0116)	(0,0107)	(0,0106)	(0,0157)	(0,0151)	(0,0148)
Profesora	-0,0246	-0,0098	-0,0123	-0,0154	-0,0056	-0,0122
Error estándar	(0,0174)	(0,0152)	(0,0150)	(0,0200)	(0,0185)	(0,0180)
Mujer × Profesora	0,0761***	0,0630***	0,0593***	0,0397*	0,0360*	0,0356*
Error estándar	(0,0144)	(0,0134)	(0,0133)	(0,0171)	(0,0165)	(0,0163)
Asociación de profesora para mujeres	0,0515**	0,0532***	0,0470**	0,0243	0,0304	0,0234
Error estándar	(0,0171)	(0,0147)	(0,0146)	(0,0187)	(0,0170)	(0,0166)
R^2 ajustado	0.491	0.536	0.540	0.421	0.445	0.448
N observaciones	73.796	73.796	73.367	73.796	73.796	73.157
RMSE	0.714	0.681	0.678	0.761	0.745	0.743
Controles incluidos						
Académicos	Sí	Sí	Sí	Sí	Sí	Sí
Familiares	No	Sí	Sí	No	Sí	Sí
Docentes	No	No	Sí	No	No	Sí

Nota: En cada modelo, la variable dependiente corresponde al puntaje SIMCE estandarizado de la asignatura analizada y la categoría de referencia corresponde a estudiantes hombres con profesor hombre. La fila “Asociación de profesora para mujeres” corresponde a la combinación lineal $\beta_2 + \beta_3$. Códigos de significancia: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.

Estos resultados permiten observar cómo se comportan los coeficientes de interés a medida que se incorporan controles adicionales.

En Matemática, el coeficiente asociado a $Mujer_i$ es negativo y estadísticamente significativo en las tres especificaciones, indicando que, cuando la asignatura es impartida por un profesor hombre, las estudiantes presentan un puntaje aproximadamente entre 0,08 y 0,09 desviaciones estándar menor que los estudiantes hombres, una vez consideradas las características incorporadas en cada modelo. Por su parte, el coeficiente de $Profesora_{is}$, que representa la asociación entre tener una profesora y el rendimiento de los estudiantes hombres, presenta valores negativos de pequeña magnitud y no resulta estadísticamente significativo en ninguna de las tres especificaciones.

En contraste, la interacción $Mujer_i \times Profesora_{is}$ es positiva y estadísticamente significativa en los tres modelos, indicando que la asociación entre tener una profesora y el rendimiento es 0,0761, 0,0630 y 0,0593 desviaciones estándar más favorable para las estudiantes que para los estudiantes hombres en los Modelos 1, 2 y 3, respectivamente. Es importante precisar que estos valores representan cuánto difiere dicha asociación entre ambos grupos y no la magnitud correspondiente específicamente a las estudiantes mujeres, por lo que, para obtener esta última, se debe considerar conjuntamente el coeficiente de $Profesora_{is}$ y la interacción $Mujer_i \times Profesora_{is}$, es decir, $\beta_2 + \beta_3$. Esta combinación alcanza 0,0515, 0,0532 y 0,0470 desviaciones estándar en los Modelos 1, 2 y 3, respectivamente, y resulta positiva y estadísticamente significativa en las tres especificaciones. Por lo tanto, además de observarse una diferencia en la relación según el género del estudiante, en Matemática se encuentra evidencia de una asociación positiva entre tener una profesora y el rendimiento de las estudiantes mujeres.

En Ciencias Naturales, los resultados presentan algunas diferencias respecto de Matemática. El coeficiente asociado a $Mujer_i$ no resulta estadísticamente significativo en ninguna de las tres espe-

cificaciones, por lo que no se encuentra evidencia de una diferencia de rendimiento entre estudiantes mujeres y hombres cuando la asignatura es impartida por un profesor hombre; de manera similar, el coeficiente de $Profesora_{is}$ tampoco alcanza significancia estadística. En cambio, la interacción $Mujer_i \times Profesora_{is}$ permanece positiva y significativa al 5 %, con valores de 0,0397, 0,0360 y 0,0356 desviaciones estándar en los Modelos 1, 2 y 3, respectivamente, mostrando que la asociación con una profesora es relativamente más favorable para las estudiantes que para los estudiantes hombres. No obstante, al considerar específicamente a las estudiantes mediante la combinación $\beta_2 + \beta_3$, las magnitudes estimadas corresponden a 0,0243, 0,0304 y 0,0234 desviaciones estándar y no resultan estadísticamente significativas al 5 %. De este modo, en Ciencias Naturales se observa una diferencia relativa en la asociación según el género del estudiante, pero no evidencia suficiente para sostener que, entre las estudiantes, tener una profesora se encuentre asociado con un rendimiento distinto al observado cuando tienen un profesor hombre.

Por otra parte, las medidas de ajuste evolucionan favorablemente a medida que se incorporan nuevos controles. En Matemática, el R^2 ajustado aumenta desde 0,491 en el Modelo 1 hasta 0,540 en el Modelo 3, mientras que el RMSE disminuye desde 0,714 hasta 0,678; en Ciencias Naturales se observa una evolución similar, con un aumento del R^2 ajustado desde 0,421 hasta 0,448 y una reducción del RMSE desde 0,761 hasta 0,743. Si bien estas variaciones son consistentes con un mayor ajuste de las especificaciones que incorporan más información, la comparación con el Modelo 3 debe considerar la leve reducción del número de observaciones asociada a la disponibilidad de las características docentes incluidas en esta especificación.

En síntesis, los resultados de ambas asignaturas muestran que la principal diferencia se encuentra en la forma en que la asociación entre el género docente y el rendimiento varía según el género del estudiante. Mientras en Matemática, además de una interacción positiva, la asociación estimada

de tener una profesora para las estudiantes resulta positiva y estadísticamente significativa en las tres especificaciones, en Ciencias Naturales se observa la diferencia relativa entre estudiantes mujeres y hombres, pero no una asociación estadísticamente significativa para las estudiantes consideradas por separado. Estos resultados constituyen una primera aproximación relevante para la investigación, aunque al basarse únicamente en características observables no permiten atribuir una interpretación causal a las relaciones estimadas, lo que motiva avanzar hacia las especificaciones más exigentes que permitan controlar fuentes adicionales de heterogeneidad no observada.

8.1.2. Resultados con efectos fijos

A partir de los resultados anteriores, el análisis avanza hacia especificaciones que incorporan efectos fijos y permiten controlar fuentes adicionales de heterogeneidad no observada. En este caso, los modelos no representan únicamente una incorporación progresiva de controles, sino que modifican la fuente de variación utilizada para estimar la relación entre el género docente y el rendimiento. Por esta razón, el Modelo 5, que incorpora efectos fijos por estudiante, constituye la especificación principal del análisis, mientras que el Modelo 4 con efectos fijos por establecimiento funciona como referencia y el Modelo 6 permite evaluar la persistencia del resultado al incorporar adicionalmente efectos fijos a nivel de curso-asignatura.

La Tabla 14 presenta los principales coeficientes de estas estimaciones, mientras que los resultados completos se encuentran en el Anexo 11.7.

Tabla 14: Resultados de los modelos con efectos fijos

	Modelo 4		Modelo 5	Modelo 6
Coefficientes	Matemática	Ciencias Naturales	Principal	Curso-asignatura
Mujer	−0,0909***	0,0063	–	–
Error estándar	(0,0077)	(0,0109)	–	–
Profesora	0,0034	0,0135	−0,0139	–
Error estándar	(0,0212)	(0,0310)	(0,0112)	–
Mujer × Profesora	0,0269**	0,0159	0,0385***	0,0249*
Error estándar	(0,0101)	(0,0130)	(0,0107)	(0,0097)
Asociación de profesora para mujeres	0,0303	0,0295	0,0245*	–
Error estándar	(0,0210)	(0,0303)	(0,0102)	–
R^2 ajustado	0.638	0.524	0.712	0.727
N observaciones	73.254	73.044	147.592	147.352
Efectos fijos incluidos				
Establecimiento	Sí	Sí	No	No
Estudiante	No	No	Sí	Sí
Asignatura	No	No	Sí	No
Curso-asignatura	No	No	No	Sí

Nota: La variable dependiente corresponde al puntaje SIMCE estandarizado. Los Modelos 5 y 6 utilizan la base en formato estudiante-asignatura, por lo que cada estudiante puede aportar una observación para Matemática y otra para Ciencias Naturales. La fila “Asociación de profesora para mujeres” corresponde a la combinación lineal $\beta_2 + \beta_3$. En el Modelo 6, el efecto principal de *Profesora* queda absorbido por los efectos fijos curso-asignatura y, por lo tanto, tampoco es posible identificar separadamente $\beta_2 + \beta_3$. El símbolo “–” indica que el coeficiente se encuentra absorbido por los efectos fijos o no puede identificarse separadamente. Códigos de significancia: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.

El resultado principal de esta etapa corresponde al Modelo 5, cuya identificación se basa en comparar el rendimiento de un mismo estudiante entre Matemática y Ciencias Naturales, aprovechando las diferencias en el género de los docentes que enfrenta en ambas asignaturas. En la muestra utilizada, 32.412 de los 73.796 estudiantes, equivalentes al 43,9 %, tienen una profesora en una de las asignaturas y un profesor hombre en la otra, lo que permite dimensionar de manera directa la proporción de estudiantes que presenta cambios en el género docente entre ambas materias.

Además de esta proporción, resulta relevante considerar cuánto de la variación total del género docente proviene de diferencias entre asignaturas para un mismo estudiante, dado que esta es

precisamente la fuente de variación utilizada por el modelo con efectos fijos individuales. A partir de la descomposición de varianza, se obtiene que el 48,8 % de la variación total del género docente corresponde a diferencias dentro del estudiante entre Matemática y Ciencias Naturales.

Bajo esta especificación, el coeficiente asociado a $Profesora_{is}$ alcanza un valor de $-0,0139$ desviaciones estándar y no resulta estadísticamente significativo, por lo que no se encuentra evidencia de una asociación entre tener una profesora y el rendimiento de los estudiantes hombres. En cambio, la interacción $Mujer_i \times Profesora_{is}$ es positiva y estadísticamente significativa, con una magnitud de $0,0385$ desviaciones estándar, indicando que la asociación entre tener una profesora y el rendimiento es más favorable para las estudiantes que para los estudiantes hombres, aun después de controlar aquellas características del estudiante que permanecen constantes entre asignaturas.

Por otro lado, al igual que en los modelos anteriores, esta interacción representa la diferencia entre ambos grupos y no la asociación correspondiente específicamente a las estudiantes mujeres. Por ello, para obtener esta última se considera conjuntamente el coeficiente de $Profesora_{is}$ y la interacción $Mujer_i \times Profesora_{is}$, cuya suma alcanza $0,0245$ desviaciones estándar y resulta positiva y estadísticamente significativa al 5 %. De este modo, el modelo principal muestra no solo que la relación con una profesora es relativamente más favorable para las estudiantes que para los estudiantes hombres, sino también que, para las estudiantes consideradas por separado, tener una profesora se encuentra asociado con un rendimiento aproximadamente $0,025$ desviaciones estándar mayor que tener un profesor hombre.

Además, como punto de referencia, el Modelo 4, al incorporar efectos fijos por colegio, muestra que en Matemática la interacción entre estudiante mujer y profesora alcanza $0,0269$ desviaciones estándar y permanece estadísticamente significativa, aunque con una magnitud menor a la observada en los modelos con controles observables. Por su parte, en Ciencias Naturales la interacción alcanza $0,0159$

desviaciones estándar y deja de ser significativa, mientras que en ambas asignaturas la asociación estimada específicamente para las estudiantes, correspondiente a $\beta_2 + \beta_3$, tampoco resulta estadísticamente significativa.

Sin embargo, la pérdida de significancia de esta asociación en el Modelo 4 no implica necesariamente que la relación desaparezca al controlar por establecimiento, ya que en Matemática, por ejemplo, la asociación estimada para las estudiantes alcanza 0,0303 desviaciones estándar, una magnitud cercana al 0,0245 obtenido en el Modelo 5; no obstante, su error estándar es considerablemente mayor (0,0210 frente a 0,0102). Por lo tanto, la diferencia en significancia puede explicarse en gran medida por una mayor precisión de la estimación en el Modelo 5.

Finalmente, el Modelo 6 incorpora simultáneamente efectos fijos por estudiante y por cada combinación curso-asignatura, permitiendo controlar tanto las características permanentes de cada alumno como aquellas condiciones comunes a los estudiantes que comparten una misma clase y asignatura. Bajo esta especificación, el efecto principal de tener una profesora queda absorbido por el efecto fijo curso-asignatura, por lo que el parámetro que continúa identificado es la interacción entre el género del estudiante y del docente. Su coeficiente alcanza 0,0249 desviaciones estándar y permanece positivo y estadísticamente significativo al 5 %, indicando que la diferencia relativa entre estudiantes mujeres y hombres persiste incluso después de controlar estas fuentes adicionales de heterogeneidad no observada, aunque presenta una magnitud menor que la obtenida en el modelo principal.

En conjunto, las especificaciones con efectos fijos refuerzan la existencia de una relación diferenciada entre el género docente y el rendimiento según el género del estudiante. En particular, el Modelo 5 muestra que la asociación con una profesora es aproximadamente 0,0385 desviaciones estándar más favorable para las estudiantes que para los estudiantes hombres, resultado que se mantiene positivo y significativo bajo la especificación más exigente del Modelo 6. Esta persistencia resulta rele-

vante, ya que indica que el patrón observado no se explica únicamente por características permanentes de los estudiantes ni por factores comunes a quienes comparten un mismo curso y asignatura.

A partir de estos resultados, se realizaron pruebas complementarias orientadas a evaluar la plausibilidad de una interpretación causal. Por una parte, al contrastar la especificación principal con una versión que excluye el rendimiento previo específico por asignatura, la interacción $Mujer_i \times Profesora_{is}$ presenta una elevada estabilidad, pasando de 0,0381 a 0,0385 desviaciones estándar una vez incorporado dicho control; por otra, las pruebas placebo muestran que el género del docente observado en 2017 no se encuentra asociado con una diferencia previa de rendimiento entre estudiantes mujeres y hombres en 2013. En conjunto, estos resultados respaldan el supuesto de identificación utilizado y permiten interpretar la interacción como un efecto causal diferencial, siempre que, una vez considerados los efectos fijos, el rendimiento previo y los demás controles incorporados, no existan factores no observados específicos de cada asignatura que determinen simultáneamente el género docente y el rendimiento (las principales pruebas utilizadas para sustentar esta interpretación se presentan en el Anexo 11.8).

En el caso de las estudiantes, la interpretación requiere algo más de cautela, ya que las pruebas placebo muestran que quienes fueron asignadas a una profesora en 2017 presentaban, cuatro años antes, un rendimiento levemente menor que aquellas que posteriormente fueron asignadas a un profesor hombre. No obstante, esta diferencia previa no parece explicar el resultado obtenido en 2017, puesto que, aun después de controlar por el rendimiento anterior, las estudiantes que tienen una profesora presentan un rendimiento mayor que aquellas que tienen un profesor hombre. De este modo, la evidencia respalda especialmente la existencia de un efecto diferencial del género docente según el género del estudiante, aunque todavía no permite establecer por sí sola cuál es el mecanismo que explica este resultado, aspecto que se analiza a continuación mediante las estrategias orientadas a contrastar las

predicciones asociadas al modelo de rol y al sesgo docente.

8.2 Análisis de los mecanismos subyacentes

Una vez identificado un efecto diferencial del género docente según el género del estudiante, el análisis se orienta a examinar qué mecanismo podría contribuir a explicar este resultado. Como se planteó en la metodología, ninguna de las estrategias utilizadas permite identificar de manera aislada y definitiva la presencia de un mecanismo determinado, por lo que su interpretación se basa en evaluar si los distintos patrones observados son consistentes con las predicciones asociadas al sesgo docente o al modelo de rol. En este sentido, más que considerar cada estimación por separado, resulta especialmente relevante analizar la evidencia de manera conjunta y observar hacia cuál de las dos explicaciones apunta con mayor consistencia.

La Tabla 15 resume los principales resultados obtenidos mediante estas estrategias, mientras que las estimaciones completas se presentan en el Anexo 11.9.

Tabla 15: Resumen de resultados asociados a los mecanismos subyacentes

Mecanismo	Resultado evaluado	Estimación	<i>p</i> -valor
Sesgo docente	Profesora $\times$ proporción de hombres en expectativas docentes	-0,251	0,223
	Profesora $\times$ proporción de hombres, probit ordenado	-0,218	0,259
	Asociación de profesora en cursos mixtos	0,018	0,084
	Asociación de profesora en cursos solo de mujeres	0,086	0,018
Modelo de rol	Asociación de profesora en Matemática	0,039	0,003
	Diferencia entre Ciencias Naturales y Matemática	-0,034	0,083
	Mujer $\times$ Profesora sobre autoconcepto matemático	0,041	0,010
	Mujer $\times$ Profesora sobre mentalidad matemática	0,029	0,053

8.2.1. Evidencia asociada al sesgo docente

La primera hipótesis evaluada corresponde al sesgo docente, según la cual el efecto del género del profesor podría surgir de diferencias en la forma en que profesores y profesoras perciben o tratan a sus estudiantes dependiendo de su género. Para examinar esta posibilidad se analizan las expectativas académicas de los docentes según la composición de género del curso y se compara el efecto asociado a tener una profesora entre cursos mixtos y cursos compuestos únicamente por mujeres.

Respecto de las expectativas, si profesores y profesoras evaluaran de manera distinta a sus estudiantes según la presencia relativa de hombres y mujeres dentro del aula, sería esperable encontrar una relación entre la composición de género del curso y las expectativas docentes que, además, variara según el género del profesor. Sin embargo, la interacción entre tener una profesora y la proporción de estudiantes hombres presenta un coeficiente negativo que no alcanza significancia estadística ($p = 0,223$), resultado que se mantiene al utilizar un modelo probit ordenado ($p = 0,259$). Por lo tanto, no se encuentra evidencia de que profesores y profesoras modifiquen de manera sistemáticamente distinta sus expectativas académicas frente a cambios en la composición de género del curso.

Esta primera evidencia se complementa con la comparación entre cursos mixtos y cursos exclusivamente femeninos, dado que, bajo la hipótesis de sesgo docente, las diferencias de trato entre estudiantes hombres y mujeres deberían adquirir especial relevancia cuando ambos grupos están presentes en una misma sala. No obstante, los resultados siguen una dirección distinta: en los cursos mixtos, la asociación estimada entre tener una profesora y el rendimiento de las estudiantes alcanza aproximadamente 0,018 desviaciones estándar y no resulta significativa al 5 %, mientras que en los cursos compuestos únicamente por mujeres esta asociación aumenta hasta cerca de 0,086 desviaciones estándar y sí resulta estadísticamente significativa.

Si bien la diferencia entre ambos tipos de curso presenta una significancia más débil y, por tanto, debe interpretarse con cautela, el patrón observado tampoco coincide con lo esperado si el efecto estuviera explicado principalmente por un trato diferencial hacia mujeres y hombres dentro de una misma aula. Así, al considerar conjuntamente las expectativas docentes y la composición de género de los cursos, la evidencia obtenida entrega un respaldo limitado a la hipótesis de sesgo docente como principal explicación del efecto encontrado.

8.2.2. Evidencia asociada al modelo de rol

Frente a la limitada evidencia encontrada para el sesgo docente, las estrategias asociadas al modelo de rol permiten analizar si la presencia de una profesora adquiere mayor relevancia en contextos donde una referente femenina podría influir con mayor fuerza sobre las estudiantes. Para ello, se consideran las diferencias entre asignaturas, el nivel educativo de la madre y distintas medidas relacionadas con las percepciones académicas de las propias estudiantes, cuyos principales resultados se resumen en la Tabla 16.

Tabla 16: Principales resultados asociados al modelo de rol

Resultado evaluado	Estimación	<i>p</i> -valor
Diferencias por asignatura		
Asociación de profesora en Matemática	0,0391**	0,003
Asociación de profesora en Ciencias Naturales	0,0055	0,713
Diferencia entre Ciencias Naturales y Matemática	-0,0336	0,083
Educación de la madre		
Asociación de profesora para estudiantes con madres de menor nivel educativo	0,0340	0,058
Diferencia: educación media respecto de educación básica o menor	-0,0044	0,813
Diferencia: educación superior respecto de educación básica o menor	-0,0212	0,333
Autoconcepto matemático		
Mujer × Profesora	0,0411*	0,010
Asociación de profesora para estudiantes mujeres	0,0664*	0,031
Mentalidad matemática		
Mujer × Profesora	0,0293	0,053
Asociación de profesora para estudiantes mujeres	0,0118	0,617
Mentalidad general		
Mujer × Profesora	-0,0098	0,525

Nota: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.

En primer lugar, la comparación entre asignaturas muestra que la asociación con una profesora se concentra principalmente en Matemática, donde alcanza 0,0391 desviaciones estándar y resulta estadísticamente significativa al 1 %, mientras que en Ciencias Naturales disminuye hasta 0,0055 desviaciones estándar y no resulta significativa. A su vez, la diferencia entre ambas asignaturas alcanza $-0,0336$ desviaciones estándar y presenta significancia al 10 %, por lo que, si bien no alcanza el nivel convencional del 5 %, entrega evidencia sugerente de que la asociación con una profesora podría ser mayor en Matemática.

Una segunda aproximación considera el nivel educativo de la madre, bajo la idea de que la relevancia de una profesora como referente podría aumentar cuando las estudiantes cuentan con menos referentes femeninos asociados a mayores niveles educativos dentro de su entorno familiar. En este caso, la asociación estimada alcanza 0,0340 desviaciones estándar entre las estudiantes cuyas madres

poseen educación básica o menor y presenta significancia al 10 %, mientras que para educación media disminuye a aproximadamente 0,0296 y para educación superior a 0,0128; sin embargo, las diferencias entre estos grupos no resultan estadísticamente significativas, por lo que el gradiente observado constituye únicamente evidencia sugerente y no permite afirmar que el efecto varíe efectivamente según la educación de la madre.

A este patrón se suman los resultados asociados a las percepciones académicas de las propias estudiantes, donde la evidencia más clara aparece en el autoconcepto matemático. En particular, la interacción entre ser estudiante mujer y tener una profesora alcanza 0,0411 desviaciones estándar y resulta estadísticamente significativa al 5 %, mientras que la asociación estimada específicamente para las estudiantes mujeres alcanza 0,0664 desviaciones estándar y también resulta significativa al 5 %; de este modo, la presencia de una profesora se relaciona no solo con un mayor rendimiento en Matemática, sino también con una percepción más favorable de las propias capacidades en esta asignatura.

En contraste, los resultados asociados a la mentalidad entregan un respaldo más débil, ya que la interacción entre ser estudiante mujer y tener una profesora sobre la mentalidad matemática alcanza 0,0293 desviaciones estándar y presenta significancia únicamente al 10 %, mientras que la asociación estimada específicamente para las estudiantes no resulta significativa. Asimismo, en la medida de mentalidad general la interacción presenta una magnitud cercana a cero y tampoco alcanza significancia estadística, por lo que la evidencia parece concentrarse principalmente en el autoconcepto matemático más que en cambios generales en las creencias de las estudiantes.

En conjunto, los resultados muestran un patrón compatible con las predicciones del modelo de rol, ya que la asociación con una profesora se concentra principalmente en Matemática, aparece acompañada por una mejora significativa en el autoconcepto matemático y presenta una tendencia coherente con una mayor relevancia de la figura docente femenina entre estudiantes cuyas madres poseen meno-

res niveles educativos. Si bien parte de esta evidencia alcanza únicamente significancia al 10 % y las diferencias asociadas a la educación materna no son concluyentes, la coincidencia entre los distintos patrones entrega un respaldo más consistente al modelo de rol que el observado para la hipótesis de sesgo docente, especialmente porque el resultado sobre autoconcepto sugiere que la presencia de una profesora puede relacionarse no solo con el desempeño de las estudiantes, sino también con la forma en que perciben sus propias capacidades.

Esta relación permite ampliar la interpretación de los resultados más allá del rendimiento académico, pues contar con referentes femeninos dentro del aula podría contribuir a que las estudiantes reconozcan con mayor confianza sus capacidades en áreas como Matemática y, al mismo tiempo, ampliar las representaciones que tienen sobre quiénes pueden desenvolverse en ellas. De esta manera, la relevancia de una profesora no se limitaría al efecto observado sobre el puntaje, sino que podría vincularse también con el cuestionamiento de estereotipos que históricamente han asociado determinadas áreas del conocimiento con los hombres y con la construcción de expectativas académicas más amplias entre las estudiantes. Aun así, estos resultados no implican que una mayor presencia de profesoras permita por sí sola reducir las desigualdades de género, sino que sugieren que la representación femenina dentro del entorno educativo puede formar parte de un conjunto más amplio de factores que contribuyen a reproducir o disminuir dichas brechas.

9 Conclusiones

La evidencia obtenida permite concluir que el emparejamiento de género entre docente y estudiante constituye una dimensión relevante para comprender las diferencias de rendimiento académico, particularmente en Matemática, donde la combinación entre estudiantes mujeres y profesoras presenta resultados más favorables. En este sentido, los hallazgos no indican que un género docente sea superior al otro en términos absolutos, sino que la relación entre el género del profesor y el desempeño depende también del género del estudiante, patrón que se mantiene al incorporar especificaciones progresivamente más exigentes y que, junto con las pruebas complementarias de identificación, respalda una interpretación causal condicionada a los supuestos planteados en esta investigación.

A partir de este resultado, el análisis de los mecanismos permite avanzar desde la identificación del efecto hacia una comprensión más amplia de los mecanismos que podrían explicar este patrón. Por una parte, la evidencia entrega un respaldo limitado a la hipótesis de sesgo docente, ya que no se observan diferencias sistemáticas en la forma en que profesores y profesoras ajustan sus expectativas frente a la composición de género del curso y tampoco aparece el patrón esperado al comparar aulas mixtas con cursos compuestos únicamente por mujeres; por otra, las estrategias asociadas al modelo de rol presentan una mayor consistencia entre sí, pues el efecto se concentra principalmente en Matemática, se acompaña de una mejora en el autoconcepto matemático de las estudiantes y muestra una tendencia compatible con una mayor relevancia de la figura docente femenina cuando existen menos referentes similares dentro del entorno familiar.

Esta última evidencia permite ampliar la interpretación del resultado más allá del rendimiento académico, ya que sugiere que la presencia de una profesora puede relacionarse también con la manera en que las estudiantes perciben sus propias capacidades. De este modo, la relevancia del emparejamiento no se limitaría únicamente al desempeño observado en una prueba, sino que podría vincularse

con procesos de identificación y representación que adquieren especial importancia en áreas como Matemática, donde continúan presentes estereotipos respecto de las capacidades de hombres y mujeres; así, contar con referentes femeninos dentro del aula podría contribuir a ampliar las posibilidades que las estudiantes reconocen para sí mismas y fortalecer su confianza para desenvolverse en espacios tradicionalmente asociados a los hombres.

Bajo esta lógica, los resultados permiten situar la representación femenina dentro de una discusión más amplia sobre las desigualdades de género presentes en la experiencia escolar, sin asumir que una mayor presencia de profesoras sea suficiente por sí sola para reducir dichas brechas, ya que estas responden a factores sociales, culturales y educativos más amplios. No obstante, la exposición cotidiana a mujeres en posiciones de conocimiento y autoridad puede formar parte de los procesos mediante los cuales estas desigualdades se reproducen o se cuestionan, especialmente cuando dicha presencia contribuye a ampliar las expectativas académicas de las estudiantes y las referencias disponibles sobre quiénes pueden desenvolverse en determinadas áreas del conocimiento.

De esta manera, el principal aporte de esta investigación consiste en actualizar y profundizar la evidencia disponible para Chile sobre el emparejamiento de género entre docentes y estudiantes, incorporando no solo una estrategia orientada a identificar su efecto sobre el rendimiento, sino también una discusión más amplia acerca de los mecanismos que podrían explicar las diferencias observadas y de las condiciones bajo las cuales esta relación adquiere mayor relevancia.

A partir de estos hallazgos, futuras investigaciones podrían profundizar el análisis extendiéndolo hacia otros niveles escolares, asignaturas y resultados posteriores, especialmente aquellos relacionados con decisiones educativas como la elección de especialidades, carreras o áreas de estudio. En esa misma línea, disponer de información que permita seguir a un mismo docente a través de distintos cursos, incorporar medidas de rendimiento previo específicas por asignatura y más cercanas al período

analizado, así como ampliar el número de materias observadas para cada estudiante, permitiría fortalecer la identificación y examinar si los patrones encontrados en Matemática se reproducen en otros ámbitos del aprendizaje; a ello se suma la posibilidad de estudiar con mayor profundidad las dimensiones socioemocionales asociadas al modelo de rol y observar de forma más directa las interacciones que se producen dentro del aula, lo que permitiría comprender con mayor precisión cuándo y para quiénes la presencia de referentes femeninos adquiere una mayor relevancia.

En conjunto, los resultados muestran que analizar el emparejamiento de género entre docente y estudiante permite comprender dimensiones del proceso educativo que van más allá de las diferencias inmediatas de rendimiento, incorporando también la forma en que las estudiantes construyen percepciones sobre sus propias capacidades y las posibilidades que reconocen como alcanzables. Así, considerar estas dinámicas permite ampliar la discusión sobre las brechas de género en educación y entender que las trayectorias académicas no se configuran únicamente a partir de lo que los estudiantes aprenden, sino también de los referentes con los que se relacionan y de las expectativas que, a partir de esas experiencias, comienzan a construir sobre sí mismos y sobre su futuro.

10 Limitaciones

Si bien la estrategia utilizada permite controlar distintas fuentes de heterogeneidad observada y no observada, la interpretación de los resultados se encuentra condicionada por algunas características de la información disponible.

En primer lugar, las bases utilizadas no cuentan con un identificador único para cada docente, lo que impide seguir a un mismo profesor entre distintos cursos e incorporar efectos fijos a este nivel, por esta razón, aunque se consideran diversas características observables del docente y se incorporan efectos fijos por curso-asignatura, no es posible aislar completamente aquellas características permanentes del profesor que podrían encontrarse relacionadas tanto con su género como con el rendimiento de los estudiantes.

A ello se suma que, si bien se dispone de resultados SIMCE previos específicos para Matemática y Ciencias Naturales, estos corresponden a cuarto básico y, por tanto, anteceden en cuatro años al período analizado. Contar con medidas de desempeño por asignatura más próximas a 2017, como las calificaciones obtenidas durante los años inmediatamente anteriores, habría permitido caracterizar con mayor precisión las diferencias iniciales de rendimiento entre materias para un mismo estudiante y fortalecer la estrategia utilizada para descartar una asignación docente relacionada con su desempeño previo. En consecuencia, aun cuando las pruebas de identificación realizadas entregan resultados favorables, la interpretación causal continúa dependiendo del supuesto de que no existan factores no observados específicos de cada asignatura que influyan simultáneamente en la asignación del género docente y en el rendimiento.

Asimismo, la estrategia con efectos fijos por estudiante se construye a partir de Matemática y Ciencias Naturales, por lo que cada alumno es observado únicamente en dos asignaturas. Esto proporciona variación suficiente para realizar la comparación dentro del estudiante, aunque limita las

posibilidades de explotar diferencias en el género docente entre un conjunto más amplio de materias y, al mismo tiempo, restringe la generalización de los resultados hacia otras áreas del aprendizaje, especialmente aquellas en las que las dinámicas de género podrían manifestarse de manera diferente.

Finalmente, las estrategias utilizadas para distinguir entre el modelo de rol y el sesgo docente permiten contrastar distintas predicciones asociadas a ambos mecanismos, pero no observarlos de manera directa. Variables como las expectativas docentes, la composición de género del curso, la educación de la madre, el autoconcepto o la mentalidad entregan aproximaciones a estas dinámicas, aunque no permiten conocer, por ejemplo, si las estudiantes identifican efectivamente a sus profesoras como referentes ni cómo se producen las interacciones de género dentro del aula. Por esta razón, la mayor consistencia encontrada con el modelo de rol debe entenderse como evidencia compatible con este mecanismo y no como una demostración de que constituya la única explicación posible del efecto observado.

En conjunto, estas consideraciones delimitan el alcance de los resultados al contexto analizado, correspondiente a estudiantes de octavo básico evaluados en 2017 y a las asignaturas de Matemática y Ciencias Naturales, sin invalidar los patrones encontrados. Más bien, permiten precisar las condiciones bajo las cuales deben interpretarse y muestran aquellos aspectos en los que una mayor disponibilidad de información podría fortalecer y ampliar la evidencia presentada.

11 Anexos

11.1 Anexo A: Preguntas utilizadas en los indicadores de mentalidad

Los indicadores de mentalidad docente y mentalidad general de los estudiantes se construyeron a partir de afirmaciones orientadas a captar creencias sobre la posibilidad de desarrollar la inteligencia y las capacidades mediante el aprendizaje. Las preguntas utilizadas se presentan a continuación.

Tabla 17: Preguntas utilizadas para los indicadores de mentalidad

Variable	Grupo	Pregunta
cprof_p37_01 / cprof_p36_01	Docente	Mi inteligencia es algo que no puedo cambiar.
cprof_p37_02 / cprof_p36_02	Docente	Ponerme desafíos no me hará más inteligente.
cprof_p37_03 / cprof_p36_03	Docente	Hay ciertas cosas que simplemente no soy capaz de aprender.
cprof_p37_04 / cprof_p36_04	Docente	Si no soy naturalmente inteligente para algo, nunca me irá bien en eso.
cest_p52_01	Estudiante	Mi inteligencia es algo que no puedo cambiar.
cest_p52_02	Estudiante	Ponerme desafíos no me hará más inteligente.
cest_p52_03	Estudiante	Hay ciertas cosas que simplemente no soy capaz de aprender.

Las respuestas válidas toman valores entre 1 y 4 y fueron invertidas mediante $x^* = 5 - x$, de modo que valores mayores representaran una mentalidad más orientada al crecimiento. Para el indicador docente se exigieron al menos 3 de las 4 respuestas válidas, mientras que para la mentalidad general del estudiante se exigieron al menos 2 de las 3 respuestas; posteriormente se calculó el promedio de los ítems disponibles y se estandarizó el resultado.

11.2 Anexo B: Pregunta utilizada para la mentalidad hacia Matemática

La medida de mentalidad hacia Matemática se construyó a partir de una única pregunta del cuestionario de estudiantes.

Tabla 18: Pregunta utilizada para la mentalidad hacia Matemática

Variable	Pregunta
cest_p52_04	Si no soy naturalmente inteligente para Matemática, nunca me irá bien en esta asignatura.

Al igual que en los indicadores generales de mentalidad, las respuestas válidas toman valores entre 1 y 4 y fueron invertidas mediante $5 - x$, tras lo cual la variable fue estandarizada para que valores mayores representaran una mayor creencia en que la capacidad matemática puede desarrollarse mediante el aprendizaje.

11.3 Anexo C: Ítems utilizados para el autoconcepto matemático

El indicador de autoconcepto matemático se construyó a partir de tres preguntas del cuestionario de estudiantes, orientadas a recoger distintas dimensiones de la percepción que cada estudiante tiene sobre sus propias capacidades en Matemática.

Tabla 19: Ítems utilizados para el autoconcepto matemático

Variable	Dimensión considerada
cest_p02_09	Capacidad percibida para aprender Matemática.
cest_p40_01	Percepción de buen rendimiento en Matemática.
cest_p40_02	Percepción de facilidad para aprender Matemática.

Las respuestas válidas corresponden a valores entre 1 y 4, mientras que las categorías 0, 99 y los valores faltantes fueron tratados como información no disponible. Dado que las preguntas se encuentran formuladas de manera que una respuesta mayor refleja una percepción más favorable de las propias capacidades, no fue necesario invertir la escala; posteriormente se calculó el promedio de los ítems disponibles exigiendo al menos 2 de las 3 respuestas válidas y el indicador resultante fue estandarizado.

11.4 Anexo D: Variables utilizadas en el índice de condiciones laborales

El índice de condiciones laborales docentes se construyó a partir de cuatro dimensiones: estabilidad contractual, adecuación de la jornada, proporción de horas no lectivas y trabajo realizado fuera de la jornada laboral.

Tabla 20: Variables utilizadas para el índice de condiciones laborales

Variable	Descripción
cprof_p08	Tipo de contrato del docente en el establecimiento.
cprof_p09_01	Total de horas semanales estipuladas en el contrato.
cprof_p09_02	Horas no lectivas semanales definidas por contrato.
cprof_p09_03	Horas semanales destinadas a actividades no lectivas fuera de la jornada laboral.

Para la estabilidad contractual se asignó un valor de 2 al contrato indefinido, 1 al contrato a plazo fijo y 0 a las restantes modalidades, mientras que la adecuación horaria se codificó con 0 para jornadas inferiores a 22 horas, 2 para jornadas entre 22 y 45 horas y 1 para jornadas superiores a 45 y de hasta 60 horas. La proporción de horas no lectivas se calculó como la razón entre horas no lectivas y horas totales contratadas, en tanto que las horas trabajadas fuera de la jornada se incorporaron con signo inverso, dado que un mayor valor representa una condición menos favorable.

Una vez contruidos, los cuatro componentes fueron estandarizados y el índice final se obtuvo como el promedio de los componentes válidos disponibles, exigiendo información para al menos 3 de las 4 dimensiones; de esta forma, valores mayores del índice representan condiciones laborales relativamente más favorables dentro de la muestra analizada.

11.5 Anexo E: Robustez a la categorización del ingreso del hogar

Se presentan las estimaciones completas del Modelo 2 utilizando los 15 tramos originales de ingreso del hogar disponibles en el cuestionario de apoderados.

Modelo 2 robustez - Matemática

```

OLS estimation, Dep. Var.: ptje_mate17_z
Observations: 73,796
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
(Intercept)    -0.579126   0.046982 -12.326681 < 2.2e-16 ***
gen_alumujer    -0.082484   0.010647  -7.747427 1.2408e-14 ***
genero_profmateprofesora -0.008746   0.015062  -0.580653 5.6151e-01
ptje_mate4b_z    0.613539   0.004066 150.891424 < 2.2e-16 ***
prom_gral_z     0.057752   0.002696  21.418722 < 2.2e-16 ***
educ_padrecon_info_padre -0.000581   0.012127  -0.047923 9.6178e-01
educ_madremedia  0.114790   0.008453  13.580387 < 2.2e-16 ***
educ_madresuperior 0.234271   0.010451  22.415129 < 2.2e-16 ***
ingreso_hogar_original100001_200000 0.045770   0.016421   2.787345 5.3448e-03 **
ingreso_hogar_original200001_300000 0.082146   0.016179   5.077306 4.0414e-07 ***
ingreso_hogar_original300001_400000 0.118867   0.016582   7.168498 9.3237e-13 ***
ingreso_hogar_original400001_500000 0.166486   0.016933   9.832094 < 2.2e-16 ***
ingreso_hogar_original500001_600000 0.194355   0.017938  10.834673 < 2.2e-16 ***
ingreso_hogar_original600001_800000 0.230380   0.018220  12.644360 < 2.2e-16 ***
ingreso_hogar_original800001_1000000 0.276839   0.019026  14.550558 < 2.2e-16 ***
ingreso_hogar_original1000001_1200000 0.317855   0.021743  14.619035 < 2.2e-16 ***
ingreso_hogar_original1200001_1400000 0.364407   0.023831  15.291422 < 2.2e-16 ***
ingreso_hogar_original1400001_1600000 0.409925   0.025822  15.875202 < 2.2e-16 ***
ingreso_hogar_original1600001_1800000 0.437812   0.027752  15.775669 < 2.2e-16 ***
ingreso_hogar_original1800001_2000000 0.474512   0.028848  16.448479 < 2.2e-16 ***
ingreso_hogar_original2000001_2200000 0.525514   0.030919  16.996558 < 2.2e-16 ***
ingreso_hogar_originalmas_2200000 0.608905   0.023989  25.382312 < 2.2e-16 ***
expecta_apodcompleta_media 0.079278   0.043106   1.839141 6.5985e-02 .
expecta_apodeducacion_superior 0.283968   0.042362   6.703439 2.3868e-11 ***
gen_alumujer:genero_profmateprofesora 0.061071   0.013337   4.579218 4.8405e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.679397   Adj. R2: 0.538263

```

Modelo 2 robustez - Ciencias Naturales

OLS estimation, Dep. Var.: ptje_nat17_z

Observations: 73,796

Standard-errors: Clustered (rbd)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.417928	0.049462	-8.449536	< 2.2e-16 ***
gen_alumujer	0.017327	0.015165	1.142585	2.5329e-01
genero_profnatprofesora	-0.005444	0.018429	-0.295400	7.6771e-01
ptje_nat4b_z	0.576898	0.005157	111.872164	< 2.2e-16 ***
prom_gral_z	0.064994	0.002939	22.114679	< 2.2e-16 ***
educ_padrecon_info_padre	-0.005447	0.013211	-0.412330	6.8012e-01
educ_madremedia	0.054717	0.008486	6.448017	1.2993e-10 ***
educ_madresuperior	0.165293	0.010885	15.185645	< 2.2e-16 ***
ingreso_hogar_original100001_200000	0.002016	0.017991	0.112034	9.1080e-01
ingreso_hogar_original200001_300000	0.025843	0.017084	1.512755	1.3044e-01
ingreso_hogar_original300001_400000	0.042956	0.017859	2.405286	1.6215e-02 *
ingreso_hogar_original400001_500000	0.050264	0.018547	2.710029	6.7627e-03 **
ingreso_hogar_original500001_600000	0.067998	0.019605	3.468424	5.3025e-04 ***
ingreso_hogar_original600001_800000	0.087882	0.019760	4.447365	8.9813e-06 ***
ingreso_hogar_original800001_1000000	0.112197	0.020781	5.399133	7.1725e-08 ***
ingreso_hogar_original1000001_1200000	0.170211	0.023824	7.144543	1.1073e-12 ***
ingreso_hogar_original1200001_1400000	0.197956	0.027238	7.267705	4.5476e-13 ***
ingreso_hogar_original1400001_1600000	0.242132	0.029446	8.223026	2.8260e-16 ***
ingreso_hogar_original1600001_1800000	0.259049	0.030288	8.552726	< 2.2e-16 ***
ingreso_hogar_original1800001_2000000	0.298421	0.032188	9.271238	< 2.2e-16 ***
ingreso_hogar_original2000001_2200000	0.327154	0.033901	9.650169	< 2.2e-16 ***
ingreso_hogar_originalmas_2200000	0.440460	0.025410	17.333895	< 2.2e-16 ***
expecta_apodcompleta_media	0.121778	0.043440	2.803402	5.0864e-03 **
expecta_apodeduccion_superior	0.230720	0.042946	5.372329	8.3144e-08 ***
gen_alumujer:genero_profnatprofesora	0.035897	0.016558	2.168006	3.0230e-02 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.744548 Adj. R2: 0.44546

11.6 Anexo F: Resultados completos de los modelos con controles observables

A continuación se presentan las estimaciones completas correspondientes a los Modelos 1, 2 y 3 para Matemática y Ciencias Naturales.

Modelo 1 - Matemática

```
OLS estimation, Dep. Var.: ptje_mate17_z
Observations: 73,796
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
(Intercept)      0.037412   0.014535    2.57386 1.0101e-02 *
gen_alumujer     -0.088938   0.011568   -7.68822 1.9579e-14 ***
genero_profmateprofesora -0.024634   0.017432   -1.41314 1.5771e-01
ptje_mate4b_z     0.685231   0.004878  140.46959 < 2.2e-16 ***
prom_gral_z      0.066390   0.003096   21.44669 < 2.2e-16 ***
gen_alumujer:genero_profmateprofesora 0.076106   0.014416    5.27930 1.3811e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.713755   Adj. R2: 0.490512
```

Modelo 1 - Ciencias Naturales

```
OLS estimation, Dep. Var.: ptje_nat17_z
Observations: 73,796
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
(Intercept)     -0.015589   0.018659   -0.835451 0.403525
gen_alumujer      0.022149   0.015665    1.413863 0.157497
genero_profnatprofesora -0.015404   0.020016   -0.769589 0.441599
ptje_nat4b_z      0.635975   0.005696  111.654396 < 2.2e-16 ***
prom_gral_z      0.070559   0.003150   22.397800 < 2.2e-16 ***
gen_alumujer:genero_profnatprofesora 0.039690   0.017086    2.322948 0.020243 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.760937   Adj. R2: 0.420927
```

Modelo 2 - Matemática

```
OLS estimation, Dep. Var.: ptje_mate17_z
Observations: 73,796
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
(Intercept)    -0.523998    0.045869  -11.423845   < 2.2e-16 ***
gen_alumujer    -0.084397    0.010676   -7.905167  3.6228e-15 ***
genero_profmateprofesora -0.009750    0.015155   -0.643375  5.2003e-01
ptje_mate4b_z    0.615992    0.004089  150.651702   < 2.2e-16 ***
prom_gral_z     0.057845    0.002703   21.399786   < 2.2e-16 ***
educ_padrecon_info_padre 0.012941    0.012062    1.072927  2.8338e-01
educ_madremedia 0.131054    0.008603   15.233463   < 2.2e-16 ***
educ_madresuperior 0.269235    0.010548   25.524699   < 2.2e-16 ***
ingreso_hogar_4medio 0.140799    0.008228   17.111903   < 2.2e-16 ***
ingreso_hogar_4medio_alto 0.319796    0.016550   19.322441   < 2.2e-16 ***
ingreso_hogar_4alto 0.475168    0.018652   25.474979   < 2.2e-16 ***
expecta_apodcompleta_media 0.086849    0.043154    2.012544  4.4244e-02 *
expecta_apodeducacion_superior 0.302886    0.042375    7.147709  1.0825e-12 ***
gen_alumujer:genero_profmateprofesora 0.062961    0.013354    4.714761  2.5207e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.680837   Adj. R2: 0.536373
```

Modelo 2 - Ciencias Naturales

```
OLS estimation, Dep. Var.: ptje_nat17_z
Observations: 73,796
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
(Intercept)    -0.404960    0.047680   -8.493256   < 2.2e-16 ***
gen_alumujer    0.017141    0.015139    1.132271  2.5760e-01
genero_profnatprofesora -0.005584    0.018465   -0.302409  7.6236e-01
ptje_nat4b_z    0.578555    0.005177  111.763849   < 2.2e-16 ***
prom_gral_z     0.065155    0.002940   22.159482   < 2.2e-16 ***
educ_padrecon_info_padre 0.000713    0.013144    0.054241  9.5675e-01
educ_madremedia 0.060580    0.008517    7.112486  1.3927e-12 ***
educ_madresuperior 0.183143    0.010816   16.932244   < 2.2e-16 ***
ingreso_hogar_4medio 0.069897    0.008240    8.482855   < 2.2e-16 ***
ingreso_hogar_4medio_alto 0.223293    0.017787   12.553874   < 2.2e-16 ***
ingreso_hogar_4alto 0.380590    0.018798   20.246466   < 2.2e-16 ***
expecta_apodcompleta_media 0.123935    0.043358    2.858397  4.2847e-03 **
expecta_apodeducacion_superior 0.236979    0.042890    5.525325  3.5453e-08 ***
gen_alumujer:genero_profnatprofesora 0.036029    0.016536    2.178756  2.9421e-02 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

RMSE: 0.745085 Adj. R2: 0.444742

Modelo 3 - Matemática

OLS estimation, Dep. Var.: ptje_matel7_z

Observations: 73,367

Standard-errors: Clustered (rbd)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.610405	0.160109	-3.812434	1.4015e-04 ***
gen_alumujer	-0.084164	0.010576	-7.958005	2.3914e-15 ***
genero_profmateprofesora	-0.012304	0.014969	-0.821930	4.1118e-01
ptje_mate4b_z	0.612058	0.004092	149.582364	< 2.2e-16 ***
prom_gral_z	0.058396	0.002717	21.492956	< 2.2e-16 ***
educ_padrecon_info_padre	0.015101	0.011998	1.258569	2.0828e-01
educ_madremedia	0.125476	0.008576	14.631201	< 2.2e-16 ***
educ_madresuperior	0.256871	0.010577	24.285558	< 2.2e-16 ***
ingreso_hogar_4medio	0.134341	0.008194	16.395841	< 2.2e-16 ***
ingreso_hogar_4medio_alto	0.309292	0.016663	18.561167	< 2.2e-16 ***
ingreso_hogar_4alto	0.450952	0.019032	23.694950	< 2.2e-16 ***
expecta_apodcompleta_media	0.098420	0.043538	2.260569	2.3852e-02 *
expecta_apodeducacion_superior	0.307165	0.042760	7.183408	8.3871e-13 ***
titulo_profmate_ftitulado	0.016428	0.153714	0.106874	9.1490e-01
especializacion_profmate_fespecializado	0.127518	0.015386	8.287702	< 2.2e-16 ***
experiencia_mate_z	0.048054	0.009989	4.810804	1.5714e-06 ***
experiencia_mate_z2	-0.025160	0.008237	-3.054621	2.2717e-03 **
INDCOND_MATE	0.018725	0.015033	1.245574	2.1301e-01
INMIND_MATE	0.014962	0.006641	2.253040	2.4323e-02 *
gen_alumujer:genero_profmateprofesora	0.059341	0.013320	4.455171	8.6645e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.678328 Adj. R2: 0.539597

Modelo 3 - Ciencias Naturales

OLS estimation, Dep. Var.: ptje_nat17_z

Observations: 73,157

Standard-errors: Clustered (rbd)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.422404	0.088717	-4.761224	2.0081e-06 ***
gen_alumujer	0.014823	0.014825	0.999882	3.1744e-01
genero_profnatprofesora	-0.012223	0.017969	-0.680215	4.9642e-01
ptje_nat4b_z	0.572650	0.005042	113.571624	< 2.2e-16 ***
prom_gral_z	0.066235	0.002930	22.606560	< 2.2e-16 ***

```
educ_padrecon_info_padre      -0.000682   0.013138  -0.051888  9.5862e-01
educ_madremedia               0.047619   0.008499   5.602659  2.2864e-08 ***
educ_madresuperior            0.160468   0.010991  14.600024  < 2.2e-16 ***
ingreso_hogar_4medio          0.057034   0.008181   6.971800  3.7719e-12 ***
ingreso_hogar_4medio_alto     0.202101   0.017668  11.438799  < 2.2e-16 ***
ingreso_hogar_4alto           0.352829   0.019230  18.348234  < 2.2e-16 ***
expecta_apodcompleta_media    0.125994   0.043421   2.901700  3.7363e-03 **
expecta_apodeducacion_superior 0.230091   0.042932   5.359484  8.9268e-08 ***
titulo_profnat_ftitulado      -0.026339   0.075533  -0.348701  7.2734e-01
especializacion_profnat_fespecializado 0.118292   0.013757   8.598669  < 2.2e-16 ***
experiencia_nat_z             0.004633   0.010104   0.458556  6.4658e-01
experiencia_nat_z2            -0.010734   0.006544  -1.640232  1.0105e-01
INDCOND_NAT                   0.044603   0.014369   3.104212  1.9242e-03 **
INMIND_NAT                    0.035024   0.006122   5.720780  1.1563e-08 ***
gen_alumujer:genero_profnatprofesora 0.035628   0.016260   2.191106  2.8515e-02 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.742674   Adj. R2: 0.448402
```

Asociación de profesora para estudiantes mujeres - M1 a M3

	Modelo	Asignatura	Estimacion	Error_estandar	t	p_value
genero_profmateprofesora	Modelo 1	Matemática	0.05147170	0.01706371	3.016442	0.0025575975
genero_profnatprofesora	Modelo 1	Ciencias Naturales	0.02428654	0.01870088	1.298684	0.1940524221
genero_profmateprofesora1	Modelo 2	Matemática	0.05321082	0.01472480	3.613688	0.0003018724
genero_profnatprofesora1	Modelo 2	Ciencias Naturales	0.03044493	0.01698396	1.792569	0.0730418787
genero_profmateprofesora2	Modelo 3	Matemática	0.04703742	0.01461003	3.219529	0.0012840123
genero_profnatprofesora2	Modelo 3	Ciencias Naturales	0.02340510	0.01660522	1.409503	0.1586864722

11.7 Anexo G: Resultados completos de los modelos con efectos fijos

A continuación se presentan las estimaciones completas correspondientes a los Modelos 4, 5 y

6.

Modelo 4 - Matemática

```
OLS estimation, Dep. Var.: ptje_mate17_z
Observations: 73,254
Fixed-effects: rbd: 3,148
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
gen_alumujer      -0.090930   0.007691 -11.822498 < 2.2e-16 ***
genero_profmateprofesora    0.003388   0.021159   0.160113 8.7280e-01
ptje_mate4b_z       0.553337   0.003207 172.556825 < 2.2e-16 ***
prom_gral_z        0.063506   0.002404  26.412799 < 2.2e-16 ***
educ_padrecon_info_padre    -0.011956   0.010976  -1.089309 2.7610e-01
educ_madremedia      0.045861   0.006965   6.584140 5.3408e-11 ***
educ_madresuperior    0.085713   0.008542  10.034080 < 2.2e-16 ***
ingreso_hogar_4medio    0.034121   0.006020   5.667916 1.5760e-08 ***
ingreso_hogar_4medio_alto  0.095017   0.013104   7.250767 5.1918e-13 ***
ingreso_hogar_4alto      0.120045   0.013107   9.158941 < 2.2e-16 ***
expecta_apodcompleta_media  0.090419   0.040275   2.245028 2.4835e-02 *
expecta_apodeducacion_superior 0.226128   0.039610   5.708852 1.2431e-08 ***
titulo_profmate_ftitulado   0.043682   0.056931   0.767267 4.4298e-01
especializacion_profmate_fespecializado 0.005955   0.036905   0.161356 8.7182e-01
experiencia_mate_z        0.013616   0.014600   0.932643 3.5108e-01
experiencia_mate_z2       -0.014078   0.010184  -1.382444 1.6693e-01
INDCOND_MATE           0.045909   0.021037   2.182280 2.9163e-02 *
INMIND_MATE            0.005218   0.009847   0.529913 5.9621e-01
gen_alumujer:genero_profmateprofesora 0.026904   0.010094   2.665476 7.7272e-03 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.588304      Adj. R2: 0.638096
Within R2: 0.418504
```

Modelo 4 - Ciencias Naturales

```
OLS estimation, Dep. Var.: ptje_nat17_z
Observations: 73,044
Fixed-effects: rbd: 3,146
```

Standard-errors: Clustered (rbd)

	Estimate	Std. Error	t value	Pr(> t)
gen_alumujer	0.006255	0.010919	0.572916	5.6674e-01
genero_profnatprofesora	0.013548	0.030959	0.437623	6.6169e-01
ptje_nat4b_z	0.533703	0.003854	138.468228	< 2.2e-16 ***
prom_gral_z	0.067464	0.002708	24.915117	< 2.2e-16 ***
educ_padrecon_info_padre	-0.015207	0.012503	-1.216212	2.2400e-01
educ_madremedia	0.031108	0.007461	4.169585	3.1339e-05 ***
educ_madresuperior	0.080437	0.009486	8.479579	< 2.2e-16 ***
ingreso_hogar_4medio	0.005882	0.006848	0.858927	3.9045e-01
ingreso_hogar_4medio_alto	0.052318	0.014557	3.593985	3.3065e-04 ***
ingreso_hogar_4alto	0.083646	0.016536	5.058462	4.4689e-07 ***
expecta_apodcompleta_media	0.125719	0.042576	2.952797	3.1724e-03 **
expecta_apodeducacion_superior	0.196452	0.041819	4.697692	2.7440e-06 ***
titulo_profnat_ftitulado	0.125136	0.056915	2.198645	2.7976e-02 *
especializacion_profnat_fespecializado	0.034602	0.040636	0.851522	3.9454e-01
experiencia_nat_z	-0.004313	0.018638	-0.231401	8.1702e-01
experiencia_nat_z2	-0.002584	0.013973	-0.184911	8.5331e-01
INDCOND_NAT	-0.009807	0.030500	-0.321539	7.4782e-01
INMIND_NAT	0.019097	0.015749	1.212586	2.2538e-01
gen_alumujer:genero_profnatprofesora	0.015931	0.012980	1.227401	2.1976e-01

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.674795 Adj. R2: 0.524242

Within R2: 0.33073

Modelo 5 - Efectos fijos por estudiante

OLS estimation, Dep. Var.: puntaje_z

Observations: 147,592

Fixed-effects: idalumno: 73,796, asignatura: 2

Standard-errors: Clustered (rbd)

	Estimate	Std. Error	t value	Pr(> t)
profesora	-0.013941	0.011163	-1.248825	0.21181795
mujer_profesora	0.038456	0.010662	3.606960	0.00031442 ***
puntaje_previo_z	0.231184	0.006023	38.384083	< 2.2e-16 ***
mujer_puntaje_previo	0.004769	0.007784	0.612668	0.54013815
titulo_proftitulado	0.181350	0.092454	1.961526	0.04990207 *
especializacion_profespecializado	-0.030198	0.010775	-2.802715	0.00509717 **
experiencia_z	0.004687	0.005801	0.807898	0.41920769
experiencia_z2	-0.008091	0.004171	-1.939829	0.05248599 .
mujer_naturales	0.056033	0.006704	8.358334	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.379562 Adj. R2: 0.711824

Within R2: 0.052047

Modelo 6 - Efectos fijos por estudiante y curso-asignatura

```
OLS estimation, Dep. Var.: puntaje_z
Observations: 147,352
Fixed-effects: idalumno: 73,676, id_curso_asignatura: 8,810
Standard-errors: Clustered (rbd)

      Estimate Std. Error   t value   Pr(>|t|)
mujer_profesora      0.024860   0.009714   2.559233   0.010536 *
puntaje_previo_z      0.237858   0.005699  41.733745 < 2.2e-16 ***
mujer_puntaje_previo -0.005540   0.007554  -0.733368   0.463388
mujer_naturales       0.065394   0.006082  10.751892 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.346902   Adj. R2: 0.726625

Within R2: 0.055332
```

Asociación de profesora para estudiantes mujeres - M4 y M5

	Modelo	Asignatura	Estimacion	Error_estandar	t	p_value
genero_profmateprofesora	Modelo 4	Matemática	0.03029225	0.02097601	1.444138	0.14870009
genero_profnatprofesora	Modelo 4	Ciencias Naturales	0.02947951	0.03027352	0.973772	0.33016976
profesora	Modelo 5	Ambas	0.02451472	0.01020536	2.402143	0.01629935

11.8 Anexo H: Pruebas complementarias de identificación

Este anexo presenta las principales pruebas utilizadas para evaluar la plausibilidad de la estrategia de identificación y la estabilidad del resultado principal.

Tabla 21: Síntesis de pruebas complementarias de identificación

Prueba	Resultado principal	<i>p</i> -valor
Asignación: rendimiento previo	-0,0116	0,051
Asignación: mujer $\times$ rendimiento previo	-0,0013	0,841
Placebo M5: mujer $\times$ profesora	-0,0012	0,892
Placebo M6: mujer $\times$ profesora	0,0024	0,791
M5 sin rendimiento previo: mujer $\times$ profesora	0,0381	< 0,001
M5 principal: mujer $\times$ profesora	0,0385	< 0,001
Balance conjunto en Matemática	Wald = 2,218	0,011
Balance conjunto en Ciencias Naturales	Wald = 0,557	0,865

En la muestra analizada, 32.412 estudiantes, equivalentes al 43,9 %, presentan un cambio en el género docente entre Matemática y Ciencias Naturales; adicionalmente, el 48,8 % de la variación total del género docente corresponde a diferencias dentro del estudiante entre ambas asignaturas.

Asignación del género docente y rendimiento previo

```

OLS estimation, Dep. Var.: profesora
Observations: 147,592
Fixed-effects: idalumno: 73,796, asignatura: 2
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value Pr(>|t|)
puntaje_previo_z    -0.011569   0.005917 -1.955139 0.050651 .
mujer_puntaje_previo -0.001300   0.006487 -0.200390 0.841188
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.321356   Adj. R2: 0.082419
Within R2: 2.039e-4

```

Prueba placebo equivalente al Modelo 5

```

OLS estimation, Dep. Var.: puntaje_previo_z

```

```
Observations: 147,592
Fixed-effects: idalumno: 73,796, asignatura: 2
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
profesora      -0.017083   0.008052  -2.121541  0.033951 *
mujer_profesora -0.001201   0.008810  -0.136304  0.891589
titulo_proftitulado  0.019431   0.071347   0.272344  0.785374
especializacion_profespecializado 0.003726   0.008469   0.439969  0.659988
experiencia_z     0.000617   0.004246   0.145359  0.884436
experiencia_z2    -0.001686   0.003019  -0.558567  0.576495
mujer_naturales   -0.094579   0.006030 -15.686004 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.373932    Adj. R2: 0.720317
                Within R2: 0.004214
```

Prueba placebo equivalente al Modelo 6

```
OLS estimation, Dep. Var.: puntaje_previo_z
Observations: 147,352
Fixed-effects: idalumno: 73,676, id_curso_asignatura: 8,810
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
mujer_profesora  0.002388   0.009011   0.265016  0.79101
mujer_naturales -0.090239   0.006103 -14.785119 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.353679    Adj. R2: 0.715723
                Within R2: 0.003408
```

Modelo 5 sin rendimiento previo

```
OLS estimation, Dep. Var.: puntaje_z
Observations: 147,592
Fixed-effects: idalumno: 73,796, asignatura: 2
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
profesora      -0.017892   0.011277  -1.586619 1.1270e-01
mujer_profesora  0.038094   0.010850  3.510977 4.5247e-04 ***
titulo_proftitulado  0.185815   0.096875  1.918095 5.5186e-02 .
especializacion_profespecializado -0.029331   0.011117  -2.638279 8.3722e-03 **
experiencia_z     0.004827   0.005913   0.816289 4.1439e-01
experiencia_z2    -0.008487   0.004287  -1.979904 4.7798e-02 *
```



```
mujer_naturales      0.033966   0.006895   4.925910 8.8119e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.389486      Adj. R2: 0.696567

Within R2: 0.001831
```

Asociación placebo de profesora para estudiantes mujeres

	Estimacion	Error_estandar	t	p_value
profesora	-0.01828415	0.007785832	-2.348387	0.01885492

11.9 Anexo I: Resultados completos asociados a los mecanismos subyacentes

A continuación se presentan las estimaciones completas utilizadas para contrastar las predicciones asociadas al sesgo docente y al modelo de rol.

Expectativas docentes - modelo lineal

```
OLS estimation, Dep. Var.: expectativa
Observations: 7,970
Standard-errors: Clustered (rbd)

              Estimate Std. Error  t value  Pr(>|t|)
(Intercept)    4.154363    0.085965  48.32609 < 2.2e-16 ***
profesora       0.160023    0.105062   1.52314  0.127834
prop_hombres   -0.323995    0.168970  -1.91747  0.055277 .
media_previo    1.052379    0.027252  38.61662 < 2.2e-16 ***
profesora:prop_hombres -0.250862    0.206017  -1.21767  0.223447
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 1.13475  Adj. R2: 0.214988
```

Expectativas docentes - probit ordenado

```
t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)
profesora     0.128355    0.099813   1.2860  0.19850
prop_hombres  -0.309067    0.158183  -1.9539  0.05075 .
media_previo   1.034100    0.033510  30.8593 < 2e-16 ***
profesora:prop_hombres -0.217756    0.192693  -1.1301  0.25848
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Modelo de rol según asignatura

```
OLS estimation, Dep. Var.: puntaje_z
Observations: 77,104
Fixed-effects: idalumno: 38,552, asignatura: 2
Standard-errors: Clustered (rbd)

              Estimate Std. Error  t value  Pr(>|t|)
```

```
profesora                0.039147   0.013275   2.949007  0.0032117 **
puntaje_previo_z         0.235848   0.005568  42.359714 < 2.2e-16 ***
titulo_proftitulado      0.167283   0.100195   1.669586  0.0951017 .
especializacion_profespecializado -0.021935   0.012178  -1.801277  0.0717559 .
experiencia_z            0.005234   0.006398   0.818138  0.4133412
experiencia_z2           -0.007471   0.004905  -1.523156  0.1278211
profesora:naturales      -0.033637   0.019368  -1.736738  0.0825323 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.36958      Adj. R2: 0.716484

Within R2: 0.053792
```

Robustez por asignatura

```
OLS estimation, Dep. Var.: puntaje_z
Observations: 147,352
Fixed-effects: idalumno: 73,676, id_curso_asignatura: 8,810
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
mujer_profesora      0.023253   0.012234   1.900732  5.7428e-02 .
puntaje_previo_z      0.237858   0.005699  41.734093 < 2.2e-16 ***
mujer_puntaje_previo  -0.005538   0.007554  -0.733071  4.6357e-01
mujer_naturales       0.062956   0.013160   4.784019  1.7968e-06 ***
mujer_profesora:naturales 0.003661   0.016971   0.215700  8.2924e-01
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.346902      Adj. R2: 0.726621

Within R2: 0.055332
```

Educación materna - modelo con interacciones

```
OLS estimation, Dep. Var.: puntaje_z
Observations: 77,104
Fixed-effects: idalumno: 38,552, asignatura: 2
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
profesora            0.034001   0.017958   1.893392  0.058398 .
educ_madre::2:profesora -0.004407   0.018633  -0.236525  0.813041
educ_madre::3:profesora -0.021247   0.021945  -0.968212  0.333014
puntaje_previo_z      0.236124   0.005566  42.421543 < 2.2e-16 ***
titulo_proftitulado    0.173673   0.101259   1.715131  0.086421 .
especializacion_profespecializado -0.022220   0.012170  -1.825880  0.067964 .
experiencia_z         0.004784   0.006398   0.747738  0.454675
```

```
experiencia_z2          -0.007444    0.004890 -1.522390  0.128013
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.369612      Adj. R2: 0.716427

Within R2: 0.053627
```

Composición de género del curso

```
OLS estimation, Dep. Var.: puntaje_z
Observations: 77,104
Fixed-effects: idalumno: 38,552, asignatura: 2
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
profesora          0.018144    0.010483   1.730819   0.083583 .
puntaje_previo_z    0.235889    0.005561  42.415153 < 2.2e-16 ***
titulo_proftitulado  0.174383    0.100285   1.738882   0.082154 .
especializacion_profespecializado -0.022310    0.012215  -1.826466   0.067876 .
experiencia_z       0.004635    0.006380   0.726597   0.467527
experiencia_z2      -0.007368    0.004889  -1.507070   0.131894
profesora:solo_mujeres 0.067991    0.037633   1.806690   0.070907 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.369567      Adj. R2: 0.716504

Within R2: 0.053858
```

Efecto de profesora según composición del curso

	tipo_curso	efecto_profesora	error_estandar	t	p_value
1	Curso mixto	0.01814365	0.01048269	1.730819	0.08348401
2	Curso solo de mujeres	0.08613426	0.03630370	2.372603	0.01766326

Mentalidad matemática

```
OLS estimation, Dep. Var.: INMIND_MATE_CEST
Observations: 73,682
Fixed-effects: rbd: 3,170
Standard-errors: Clustered (rbd)

              Estimate Std. Error   t value   Pr(>|t|)
gen_alumujer      -0.057740    0.011268  -5.124191  3.1671e-07 ***
genero_profmateprofesora -0.017461    0.023510  -0.742676  4.5773e-01
ptje_mate4b_z      0.247502    0.004195  58.995163 < 2.2e-16 ***
```

```
prom_gral_z                0.054916    0.003750 14.644993 < 2.2e-16 ***
titulo_profmate_ftitulado  -0.477373    0.055769 -8.559833 < 2.2e-16 ***
especializacion_profmate_fespecializado 0.047819    0.040627  1.177008 2.3928e-01
experiencia_mate_z         0.011562    0.017978  0.643117 5.2019e-01
experiencia_mate_z2        -0.010943    0.012575 -0.870283 3.8421e-01
gen_alumujer:genero_profmateprofesora  0.029292    0.015118  1.937571 5.2764e-02 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RMSE: 0.914422      Adj. R2: 0.098545

Within R2: 0.058504
```

Mentalidad general

OLS estimation, Dep. Var.: INMIND_GENERAL_CEST

Observations: 73,682

Fixed-effects: rbd: 3,170

Standard-errors: Clustered (rbd)

	Estimate	Std. Error	t value	Pr(> t)
gen_alumujer	-0.036266	0.011612	-3.123049	0.0018060 **
genero_profmateprofesora	-0.020443	0.023743	-0.860981	0.3893139
ptje_mate4b_z	0.175902	0.004342	40.507787	< 2.2e-16 ***
prom_gral_z	0.031930	0.003657	8.730436	< 2.2e-16 ***
titulo_profmate_ftitulado	0.163383	0.057889	2.822356	0.0047969 **
especializacion_profmate_fespecializado	0.032901	0.045398	0.724713	0.4686813
experiencia_mate_z	0.018194	0.017002	1.070136	0.2846396
experiencia_mate_z2	-0.015237	0.011147	-1.366988	0.1717260
gen_alumujer:genero_profmateprofesora	-0.009771	0.015362	-0.636045	0.5247929

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.933068 Adj. R2: 0.074224
Within R2: 0.028766

Autoconcepto matemático

OLS estimation, Dep. Var.: autoeficacia_mate_z

Observations: 72,885

Fixed-effects: rbd: 3,167

Standard-errors: Clustered (rbd)

	Estimate	Std. Error	t value	Pr(> t)
gen_alumujer	-0.211252	0.012089	-17.474317	< 2.2e-16 ***
genero_profmateprofesora	0.025314	0.031662	0.799501	4.2406e-01
ptje_mate4b_z	0.360244	0.004535	79.432089	< 2.2e-16 ***
prom_gral_z	0.080267	0.003583	22.399992	< 2.2e-16 ***



```
titulo_profmate_ftitulado      -0.362628   0.079664  -4.551969  5.5147e-06 ***
especializacion_profmate_fespecializado  0.061170   0.057036   1.072476  2.8359e-01
experiencia_mate_z             0.009737   0.024292   0.400808  6.8859e-01
experiencia_mate_z2           -0.016945   0.015312  -1.106703  2.6851e-01
gen_alumujer:genero_profmateprofesora   0.041111   0.015959   2.575985  1.0040e-02 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.883979      Adj. R2: 0.183086
Within R2: 0.131642
```

Referencias

- Agencia de Calidad de la Educación (2016). Buenas prácticas en la reducción de las brechas de género en resultados simce de comprensión de lectura y matemática ii medio. Technical report, Agencia de Calidad de la Educación, Santiago de Chile. https://archivos.agenciaeducacion.cl/Estudio_Brechas_de_Genero.pdf.
- Agencia de Calidad de la Educación (s.f.). Simce. <https://www.agenciaeducacion.cl/evaluar/simce/>.
- Angrist J. D. y Pischke J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, Princeton. pp. 235, 246.
- Carlana M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224. <https://doi.org/10.1093/qje/qjz008>.
- Centro de Estudios M. d. E. d. C. (2024). Brechas de género en la educación superior: Participación, trayectorias y resultados serie evidencias n.º 64. https://educacionsinbrechas.mineduc.cl/wp-content/uploads/sites/129/2025/03/EVIDENCIAS-64_2024-CEM.pdf.
- Cho I. (2012). The effect of teacher–student gender matching: Evidence from oecd countries. *Economics of Education Review*, 31(1):54–67. <https://doi.org/10.1016/j.econedurev.2012.02.002>.
- Coleman J. S. (1968). Equality of educational opportunity. *Equity & Excellence in Education*, 6(5):19–28. <https://doi.org/10.1080/0020486680060504>.

- Dee T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human resources*, 42(3):528–554. <https://jhr.uwpress.org/content/42/3/528>.
- Gujarati D., Guerrero D., y Medina G. (2004). *Econometría*. McGraw-Hill. <https://books.google.cl/books?id=8RttQgAACAAJ>.
- Gustafsson J.-E., Nilsen T., y Hansen K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. evidence from 50 countries in timss 2011. *Studies in Educational Evaluation*, 57:16–30. <https://doi.org/10.1016/j.stueduc.2016.09.004>.
- Lim J. y Meer J. (2017). The impact of teacher–student gender matches: Random assignment evidence from south korea. *Journal of Human Resources*, 52(4):979–997. <https://doi.org/10.3368/jhr.52.4.1215-7585R1>.
- Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (2025). Cuarta radiografía de género en ciencia, tecnología, conocimiento e innovación. https://www.minciencia.gob.cl/uploads/filer_public/b0/c8/b0c8641b-6df9-4556-bbb5-5b3691b0bfbf/4taradiografia_oficial_2025.pdf.
- Ministerio de Educación de Chile (2009). Ley General de Educación (LGE). <https://www.bcn.cl/leychile/navegar?idNorma=1006043>.
- Ministerio de Educación de Chile (2015). Ley de Inclusión Escolar. <https://www.bcn.cl/leychile/navegar?idNorma=1078172>.
- Ministerio de Educación de Chile (2018). ¿qué es el sistema de admisión escolar

(sae)? https://www.mineduc.cl/wp-content/uploads/sites/19/2018/07/PRESENTACION_SAE.pdf. Presentación oficial del Ministerio de Educación de Chile.

Miralles-Cardona C., Cardona-Moltó M.-C., y Chiner E. (2020). La perspectiva de género en la formación inicial docente: estudio descriptivo de las percepciones del alumnado. *Educación XXI*, 23(2):231–257. <https://doi.org/10.5944/educxx1.23899>.

Morales Inga S. y Morales Tristán O. (2020). ¿por qué hay pocas mujeres científicas? una revisión de literatura sobre la brecha de género en carreras STEM. *aDResearch ESIC International Journal of Communication Research*, 22:118–133. <https://doi.org/10.7263/adresic-022-06>.

Morgenroth T., Ryan M. K., y Peters K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology*, 19(4):465–483. <https://doi.org/10.1037/gpr0000059>.

Paredes V. (2014). A teacher like me or a student like me? role model versus teacher bias effect. *Economics of Education Review*, 39:38–49. <https://doi.org/10.1016/j.econedurev.2013.12.001>.

Rodríguez-Olay L. (2023). Impulsar la igualdad en las aulas. percepción del profesorado de educación primaria sobre cuestiones de género. *Ogigia. Revista electrónica de estudios hispánicos*, 33:167–192. <https://doi.org/10.24197/ogigia.33.2023.167-192>.

Scott J. W. (1986). Gender: A useful category of historical analysis. *The American Historical Review*, 91(5):1053–1075. <https://doi.org/10.1086/ahr/91.5.1053>.

Spencer S. J., Steele C. M., y Quinn D. M. (1999). Stereotype threat and women's math performan-

ce. *Journal of Experimental Social Psychology*, 35(1):4–28. <https://doi.org/10.1006/jesp.1998.1373>.

Valiente C., Parker J. H., Swanson J., Bradley R. H., y Groh B. M. (2019). Early elementary student-teacher relationship trajectories predict girls' math and boys' reading achievement. *Early Childhood Research Quarterly*, 49:109–121. <https://doi.org/10.1016/j.ecresq.2019.05.001>.

Wooldridge J. M. (2019). *Introductory econometrics: A modern approach*. Cengage Learning, 7 edición. p. 35.